

Automated evaluation of graphical user interface metrics

Évaluateur automatique de la qualité esthétique d'une
interface graphique

Année académique 2010-2011

UCL-EPL-DÉPARTEMENT D'INGÉNIERIE INFORMATIQUE

Patrick Mbenza Buanga

UNIVERSITE CATHOLIQUE DE LOUVAIN

ECOLE POLYTECHNIQUE DE LOUVAIN

DÉPARTEMENT D'INGÉNIERIE INFORMATIQUE



Automated evaluation of graphical user interface metrics
*Evaluateur automatique de la qualité esthétique d'une
interface graphique*

Promoteur : Jean Vanderdonckt

Mémoire présenté en vue
de l'obtention du grade
de Maître en Informatique
par
Mbenza-Buanga Patrick

Louvain-la-Neuve
Année académique 2010-2011

Acknowledgements

First of all, I would like to thank the Lord Jesus-Christ for enabling me to achieve all that I have achieved. Then, I am especially thankful to my supervisor, Prof. Jean Vanderdonckt for his support and guidance; not only did he propose a very interesting subject but he also provided me with abundant documentation. I would like to thank my brother Christian, my friend David, and my mother for their moral, financial and technical support throughout this academic year.

I am also very grateful to Councilor Ms. Chantal Poncin of the INGI department, EPL, UCL, for her precious guidance throughout the year.

Of course, I thank all the volunteers who participated in my test for their patience and availability.

And finally, last but not least, I am thankful to the discoverer of caffeine, without which this thesis could never have been completed.

Table of Contents

Chapter I.	Introduction.....	8
I.1	Defining Aesthetics.....	8
I.2	Aesthetics in Human-Computer Interaction.....	8
I.3	Why does Aesthetics Matter?.....	9
I.4	Can Aesthetics be measured?.....	11
I.5	Goals.....	14
I.6	Reading plan.....	16
Chapter II.	State of the art.....	17
Chapter III.	Aesthetic Judgment and Study of the Metrics.....	23
III.1	Methodology.....	23
III.2	Aesthetic experience.....	24
III.2.1	Classifying artifacts.....	24
III.2.2	Aesthetic Judgment.....	25
III.2.3	The GUI artifact.....	29
III.3	A Brief Study of the Metrics.....	30
III.3.1	Origin of the Metrics.....	30
III.3.2	Visual design in a nutshell.....	31
III.3.3	GUI design in a nutshell.....	34
III.3.4	Limits of the metrics.....	35
III.4	Eliminating irrelevant Metrics.....	38
III.5	Selecting a minimal subset of metrics.....	39
III.5.1	Density.....	39
III.5.2	Balance.....	42
III.5.3	The discarded metrics.....	44
Chapter IV.	Implementation.....	46
IV.1	Technologies used.....	46
IV.1.1	UsiXML.....	46
IV.1.2	REST.....	46
IV.1.3	DOM.....	47
IV.1.4	Java.....	47
IV.2	Architecture.....	48
Chapter V.	Tests.....	49
V.1	Screen Selection and Automated Evaluation.....	49
V.1.1	The screens.....	49
V.1.2	Discussion.....	56
V.2	Subjective Ratings by Users.....	56
V.2.1	The experiment.....	56
V.2.2	Discussion.....	57
Conclusions.....		59
References.....		61
Appendix A: Metric Description Template.....		64
Appendix B: A study of Vanderdonckt and Gillo's « Visual Techniques ».....		65
Appendix C : Selected Metrics.....		85
Appendix D : Form used to collect users ratings.....		104
Appendix E : Typical scenario with screenshots of actual Java code.....		106

List of Figures

Figure 1: “Aesthetica” (1750) by Alexander Gottlieb Baumgarten.	8
Figure 2: Scott Adams' Dogbert's view on product design.	10
Figure 3: Apple’s iMac (in 1998) was heralded as the “aesthetic revolution in computing”.	11
Figure 4: Example of a 'bad' screen.	17
Figure 5: Output from the program for a 'bad' screen.	18
Figure 6: Example of a 'good' screen.	18
Figure 7: Output from the program for a 'good' screen.	20
Figure 8: A graphic designer's C.V.	28
Figure 9: A graphic designer's C.V.	28
Figure 10: A graphic designer's C.V.	28
Figure 11: A more classic C.V.	29
Figure 12: The psychology of color.	32
Figure 13: A website’s contact page.	36
Figure 14: A website’s contact page.	36
Figure 15: The background image influences the composition of the layout.	37
Figure 16: login page with interesting colors.	37
Figure 17: The texture of the dialog box influence our aesthetic judgment.	38
Figure 18: Contact page with perspective.	38
Figure 19: Uncluttered and cluttered layouts.	39
Figure 20: Misaligned, inconsistent and cluttered layout.	40
Figure 21: Aligned and consistent interface.	40
Figure 22: uncluttered interface.	41
Figure 23: Two versions of screens in a density study: (a) a good version, (b) a bad version.	41
Figure 24: Balanced and unbalanced layouts.	42
Figure 25: Unstable interface.	43
Figure 26: Same interface; edited in Microsoft Paint.	43
Figure 27: Structure of UsiXML; The four basic levels of the Cameleon reference framework.	46
Figure 28: XML DOM Example.	47
Figure 29: good screen that should receive good ratings from the program and the users.	49
Figure 30: UsiXML manual copy of the interface “Android”.	50
Figure 31: Unbalanced screen.	51
Figure 32: UsiXML copy of “Ubuntu firewall”.	51
Figure 33: Same interface as in fig.31 with better balance.	52
Figure 34: UsiXML copy of Ubuntu firewall.	52
Figure 35: Misaligned interface.	53
Figure 36: UsiXML copy of “Printer dialog box”.	54
Figure 37: Same interface as in fig. 35 with better alignment.	55
Figure 38: UsiXML copy of the interface in fig. 37.	55
Figure 39: Perfect unity.	86
Figure 40: Poor unity.	87
Figure 41: Horizontally and vertically symmetric layouts.	88
Figure 42: Symmetric and asymmetric layouts.	88
Figure 43: Asymmetric but stable layout.	89
Figure 44: Unstable interface.	91
Figure 45: Same interface; edited in Microsoft Paint.	91
Figure 46: This interface is too symmetric!	91
Figure 47: Unaesthetic because of bad alignment.	93
Figure 48: Same interface aligned with Microsoft Paint.	93
Figure 49: Grouped and split layouts.	94
Figure 50: Excellent grouping.	95
Figure 51: IO Repartition among the screen quadrants.	96

Figure 52: Quadrant preference.	97
Figure 53: Simple and complex layouts.....	98
Figure 54: Simple interface.....	99
Figure 55: Uncluttered and cluttered layouts.	100
Figure 56: Excellent economy.....	100
Figure 57: Proportioned and disproportioned layouts.....	102

List of Diagrams

Diagram 1: Application areas of Aesthetics.	14
Diagram 2: Aesthetic Judgment; schools of thought and available tools.	15
Diagram 3: Methodology.	23
Diagram 4: Alternative methodology.	24
Diagram 5: Typical scenario of a client requesting an assessment of a UisiXML GUI.....	48

List of Tables

Table 1: Output of the program for Android.	50
Table 2: Output of the program for Ubuntu firewall.....	51
Table 3: Output of the program for the balanced version of “Ubuntu firewall”.....	52
Table 4: Output of the program for “Printer dialog box”.....	54
Table 5: Output of the program for the aligned version of “Printer dialog box”.....	55
Table 6: Statistics on participants.....	55
Table 7: Averages of subjective ratings (low=1, medium=2, high=3).	55
Table 8: Rankings of the interfaces by the program.	55

“It is amazing how complete is the delusion that beauty is goodness.”

Leo Tolstoy

Chapter I. Introduction

I.1 Defining Aesthetics

The term “aesthetics” is derived from the Greek *αἰσθητικός* (*aisthetikos*, meaning "esthetic, sensitive, sentient"), which in turn was derived from *αἰσθάνομαι* (*aisthanomai*, meaning "I perceive, feel, sense"). It has evolved through the years, has been studied from different viewpoints, and has different meanings for different schools of thought.

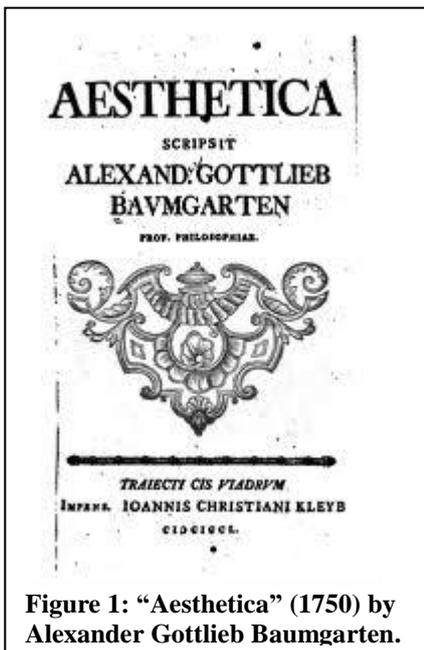


Figure 1: “Aesthetica” (1750) by Alexander Gottlieb Baumgarten.

Historically, even though, in some ancient traditions, beauty and the perception of beauty were of cosmic importance, the idea of beauty for its own sake is a relatively new concept in Western thought. Greek philosophers judged artifacts primarily on how suitable they were for their intended use and on the quality of their craftsmanship. Consequently, the idea of judging artifacts based on their creativity or the idea of art as an expression of the artist’s personality were absent from their discussions. During the renaissance, the laws of beauty were equated with the laws of nature; beautiful design had to rely on the principles of natural beauty [9].

It is not until the eighteenth century that the word “aesthetics” was introduced into philosophical terminology by Baumgarten (fig.1)[9]. It referred to the philosophical study of art and beauty. Then by the end of that century it was no longer merely a technical term in philosophy; it became an integral part of the general language.

Of the range of connotations of the term “aesthetics” that exist today in various academic disciplines and in common language, we are interested here in its meaning as “*an artistically beautiful or pleasing appearance*” (The American Heritage Dictionary of the English Language), or as “*a pleasing appearance of effect: Beauty*” (Merriam-Webster’s Collegiate Dictionary).

I.2 Aesthetics in Human-Computer Interaction

The study of human-computer interaction and user interface design is fundamentally based in quantifiable usability research. For years, many usability professionals have been planning, conducting and reporting on usability tests and their findings in order to determine how useful a system really is. Usability testing largely tests how successful users are at completing predetermined tasks or finding pieces of information with a computer system. Typically, a usability test will ask questions such as “Could you find that piece of information?” or “Were you able to

accomplish the task?” The users are observed and asked to articulate in words exactly what they are able to accomplish.[18]

So, traditionally, the human–computer interaction literature expressed only passing interest in the aesthetic aspects of the interaction. In reviewing guidelines and principles advanced by human–computer interaction experts, one cannot help but notice the marginalization of the aesthetic dimension. Discussion of aesthetic issues is infrequent and is almost invariably qualified by warnings against its potentially detrimental effects (Tractinsky, 1997). Tractinsky et al. (2000) suggest that this might be the result of two different processes. One process is characterized by a “backlash to recent tendencies by the computer industry to oversell glitz and fashion in its products...” Another process relates to the field’s origins “in disciplines that emphasize efficiency” rather than affect (p. 128).[9]

So, readers of human–computer interaction textbooks can hardly find any reference to aesthetic considerations in design. The rationales and guidelines often advocated for web site design resemble those used to direct human–computer interaction in general. While Neale and McCombe (1997) describe how to design a usable and visually appealing web site, the emphasis is clearly on the functional and the usability aspects of the design rather than on aesthetics. Similarly, Spool et al. (1999) assess web sites and draw guidelines based on usability as the sole criterion. Nielsen (2000) maintains that two basic approaches to web design exist: the ‘artistic ideal’ that reveals the designer’s self-expression and the ‘engineering ideal’ that provide solutions to users. While “there is a need for art, fun, and general good time on the web,” Nielsen contends that “the main goal of most web projects should be to make it easy for customers to perform useful tasks...” (p. 11). Hence, clear and effective communication of ideas is the design principle to follow on the Web (Lynch and Horton, 1999). [9]

What the users actually prefer is not as clear, though. Indeed, in striking contrast to the principles and the guidelines advocated by usability researchers and gurus, any random perusal of web sites would suggest that aesthetic considerations are paramount in designing for the web.

And there is a new wave of research in the visual aesthetics of computer interfaces that suggests that aesthetics is a strong determinant of pleasure experienced by the user during the interaction (Jordan, 1998) and of his satisfaction. It is also reported to be highly correlated with the system’s perceived usability (ease of use) and, to a lesser extent, with its usefulness. If once the value of computing technology was measured mostly by its usefulness for solving problems and by its usability, additional requirements, such as desirability, have now emerged and issues of visual appeal and aesthetics have become an integral part of interactive systems design.[9]

1.3 Why does Aesthetics Matter?

There is a well-established knowledge in marketing, product design, and even social psychology that beauty matters. The aesthetic quality of a product influences consumers’ attitudes, and is a major determinant of its marketplace success. Beauty can be an important quality of a product; recent research in the area of Information technology has also suggested that the visual aesthetics of user interfaces is a strong determinant of user’s satisfaction and pleasure. Furthermore, Beauty seems to influence other judgments about the same person or object. For example, not only do

people associate positive personality traits with attractive individuals, but they also tend to make more positive judgments on other attributes, such as intelligence.

Thus, several studies have suggested a correlation between the aesthetic quality of an interface and its perceived usability [7][8].(i.e. Attractive things will appear to work better). Users will commonly assign positive qualities to systems they believe are physically attractive and will report higher satisfaction with an application based on how aesthetically pleasing the system looks to them (known as the belief that “what is beautiful is usable”).

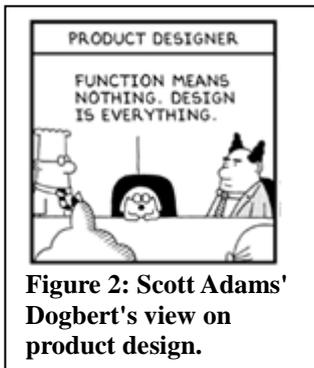


Figure 2: Scott Adams' Dogbert's view on product design.

However, the usability testing which occurs today focuses mainly on the actual usability of the system and not the role aesthetics plays into the perceived and reported usability in determining user satisfaction with an application. As mentioned earlier, usability testing tests how successful users are at completing tasks with a computer system. They are typically asked questions such as “Could you find that piece of information?” or “Were you able to accomplish the task?”

But Norman (2002) [18] says that usability barely matters to most system users. As long as the system functions at least at an acceptable level, the first comments from system users will rarely focus on ease-of-use issues.

And quite often good characteristics of an application are not even perceived by the users. They are often taken for granted and not even reported on by users because the human perceptual and attentional systems are designed to notice discrepancies, problems and distractions, not that which is expected. So, as long as the system is good enough, most people don't really care about usability. (Fig.2 humorously illustrates this tendency)

Because of this, it is important for usability professionals and testers to understand the important role aesthetics plays, to balance aesthetics and usability to ensure user interfaces are not only easy to use, but that they are also visually appealing and interesting. Thus recent research stresses the importance of studying the aesthetic aspect of user interface design and mentions the paramount need for appropriate concepts and measures of aesthetics. These could take the form of test questions or be more formally specified and implemented in a computer program.

Finally, even though some studies have contradicted the belief that “what is beautiful is usable” (like Hassenzahl(2004) and Hartmann(2008)[20] who observed that the more beautiful designs in their experiments did not have better perceived usability and concluded that aesthetics was more closely related to judgments of product “identity”) or have approached the study of aesthetics from a cultural and commercial perspective (Virginia Postrel 2003), they still agree that aesthetics is important because it becomes the dominant factor when content of competing products is similar (even leading users to forgive adverse usability experiences).

For example, since usability has become more mainstream and popular among the top websites, companies need to place more focus on aesthetic appeal to differentiate themselves from other companies and other websites [18].

Similarly, Norman (1998) suggests also that as the functionality of new IT products exceeds users' needs, and as the price of systems decrease, the competition becomes more oriented toward enhancing the users' experience rather than toward improving functionality. Once IT provides all the required features at ever-decreasing prices, considerations of (...) appearance become more important. Apple's first iMac (fig.3) is a good example of this. Norman compares this process to the state of the watch industry, which long ago passed users' technological requirements: watches are now often marketed as objects of fashion or emotion.



Figure 3: Apple's iMac (in 1998) was heralded as the "aesthetic revolution in computing".

I.4 Can Aesthetics be measured?

People are generally very skeptic as to whether Aesthetics can be measured. Conventional wisdom seems to doubt that aesthetics can be measured. It is argued that Aesthetic appeal is largely a subjective judgment that is not rooted in any preconceived notion of purpose. Beauty, they say, is largely very personal and also very cultural. Its perception depends on culture, gender, age, personality, etc.

This view matches the concepts known in philosophy as "Aesthetic Subjectivism" and "Aesthetic Relativism" which basically hold that aesthetic qualities do not exist in objects independently of subjects (i.e. people or observers) and their valuations. The aesthetic subjectivist maintains that aesthetic values, such as "beautiful," "elegant," "ugly," "sublime," or "poignant," are entirely relative to the preferences of individuals. Or, as it is popularly expressed, "beauty is in the eye of the beholder." And the relativist says that aesthetic qualities and the way they are experienced are not just dependent on single individuals but relative to culture. Where "culture" is to be understood in a broad sense. A person is influenced by the culture he or she belongs to, which means that the person's subjective reactions are not independent of it [28].

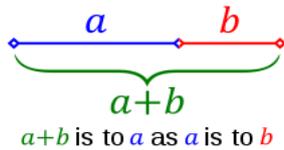
The opposite philosophical view, called "Aesthetic Objectivism", holds that the aesthetic qualities belong to objects themselves, as do other, e.g. physical qualities and are independent of the valuating subject. But since aesthetic experience depends on those aesthetic qualities, it gives us knowledge of them. Our aesthetic judgment about an object is correct and thus yields knowledge if it complies with the aesthetic qualities the object possesses [28].

This basically means that there are universal patterns of beauty – aesthetic qualities valid for humans in every culture. And all humans evaluate objects that correspond to these patterns as beautiful, provided that they can perceive these patterns.

If the objectivists are right then there must be a way of measuring beauty! Or at least some aspects of it. There must be a system of methods, principles or even formulas that could allow us to evaluate or create aesthetic qualities in objects in a systematic/scientific way.

Throughout history, many, with various levels of ambition, believed just that:

- Euclide's *Elements* (300 BC) provides the first known written definition of what is now called the golden ratio. The golden ratio, which has fascinated Western intellectuals of diverse interests for at least 2,400 years, is an irrational mathematical constant, approximately 1.61803398874989. It is generally denoted by the greek letter phi (φ). Two quantities a , b are in the golden ratio if the ratio of the sum of the quantities to the larger quantity is equal to the ratio of the larger quantity to the smaller one. Here is a geometric representation of this definition:



This geometric relationship is expressed algebraically as:

$$\frac{a+b}{a} = \frac{a}{b} \equiv \varphi.$$

This equation has one positive solution in the set of algebraic irrational numbers:

$$\varphi = \frac{1 + \sqrt{5}}{2} = 1.6180339887\dots$$

Other names frequently used for the golden ratio are the *golden section* and *golden mean*. Other terms encountered include *extreme and mean ratio*, *medial section*, *divine proportion*, *divine section*, *golden proportion*, *golden cut*, *golden number*, and *mean of Phidias*.

At least since the Renaissance, many artists and architects have proportioned their works to approximate the golden ratio—especially in the form of the golden rectangle, in which the ratio of the longer side to the shorter is the golden ratio—believing this proportion to be aesthetically pleasing.[30]

- But the first attempt to measure aesthetics was made in the early 1930s by the American mathematician George D. Birkhoff; who proposed a mathematical theory of aesthetics in his 1933 book, *Aesthetic Measure*. While writing this book, he spent a year traveling around the world studying art, music, and poetry in various countries. He then came up with a formula to measure the relative importance of aesthetic factors involved in an aesthetic experience: $M=O/C$.

The aesthetic value or measure M , according to Birkhoff, depends on the elements of Order O (symmetry in sculpture, melody in music, to name a few) and complexity C of an aesthetic object. By complexity he means that which increases a "feeling of tension" or "effort of attention" involved in the perception of an aesthetic object. According to Birkhoff, complexity negatively effects overall aesthetic measure since complex objects tend to deflect on looker's contemplation.

Birkhoff applied the formula to various polygonal forms, poetry and music, taking into consideration only the formal elements of art. This is because connotative elements "seem to defy classification since they touch our experience at so many points and in an entirely undefinable way." [31][32]

Others have followed in the footsteps of Birkhoff, especially in the last two or three decades. Naukkarinen [27] laments that "in recent years ... more and more books and articles have been published in which aesthetic issues are dealt with by methods and conceptual tools that are provided by natural sciences and other non-philosophical disciplines". Here are a few examples from the field of the human-computer interaction:

- Ngo et al. (2000, 2001) [2] [5] [6] developed fourteen aesthetic measures for graphic displays: balance, equilibrium, symmetry, sequence, cohesion, unity, proportion, simplicity, density, regularity, economy, homogeneity, rhythm, and order and complexity.
- Vanderdonckt and Gillo (1994) [17] proposed thirty "visual techniques" grouped in five categories (Physical, Composition, Association and dissociation, Ordering, Photographic techniques) to help user interface designers in arranging "information items" (e.g. text, images, diagrams, pictures, tables) in such a way that it is visually attractive, perceptive and easily understandable.

An experiment based on these visual techniques was conducted by the author in a group of five students as part of the Human-Computer Interaction course, this year. This work had multiple objectives: First, we wanted to study the visual techniques and determine whether they were good predictors of the aesthetics judgments of people; whether they captured the essence of aesthetics; if they did which ones were the best predictors of users' preferences, and which ones were redundant. Secondly, we wanted to determine whether these metrics transcended culture, age or gender. (The answers to these questions would also allow us to state our position in the "objectivism vs. subjectivism" debate based on our findings). And finally, we wanted to know if there were a correlation between perceived usability of an interface and a good appraisal of its appearance based on the thirty visual techniques.

So we selected a sample population of fourteen people (seven males and seven females) from various backgrounds, origins and ages. We selected five interfaces with different obvious characteristics and asked the users to evaluate (in a scale of 1 to 5) the metrics on the interfaces. Then we gathered the answers and conducted a statistical analysis.

The results were very divergent (we questioned our method and described its limits and the bias it might introduce) and we found no correlation between perceived usability and evaluated aesthetics; no correlation between culture, age, gender and the way the metrics are evaluated. And although some metrics seemed to be good predictors of perceived aesthetics, others seemed redundant. And we were unable to determine whether they captured the essence of good aesthetics. (See Appendix B for more details).

- Schneiderman et al. (1995,1996) [3] [4] [15] proposed thirteen metrics for checking the consistency of a graphical user interface: Area balances, Aspect Ratio, Distinct typefaces, Griddedness, Margins, Non-Widget Area, Widget density, Widget totals, Interface concordance, Button concordance, Button Layout table, Interface Speller, Terminology Basket.
- Hartmann et al. (2005) [19] and Perlman [14] also proposed metrics and tools (see section 2: State of the art).
- Balinsky[33] claimed that "symmetry is one of the most fundamental principles in design. The choice between symmetry and asymmetry affects the layout and feeling of a design. A symmetrical page gives a feeling of permanence and stability, while informal or asymmetrical balance creates interest". So she proposed and discussed a way "to solve the problem of an automatic detection of axial and radial symmetry or lack of it in published documents". "Previous approaches to this problem, she said, gave only a necessary condition for symmetry. We present a necessary and sufficient criterion for automatic symmetry detection and also introduce a Euclidean-type distance from any layout to the closest symmetrical one..."
- Vande Moere et al. [34] used a set of metrics to objectively evaluate the aesthetic quality of data visualizations.
- Tullis (1983, 1985) [14] developed six metrics for alphanumeric displays: overall density, local density, grouping, average size of groups, number of items and item uncertainty.
- Harrington et al. [35] wrote that "technology offers automated systems for document creation and a diversity of presentation methods that make the single fixed output appearance obsolete. To produce documents that look good, one needs methods to quantify document aesthetics." They described a "measure of aesthetics that should be used in automated layout of documents.

Note that there have been attempts to measure beauty in other fields: For instance, many scientists have been researching facial beauty and human attractiveness in recent decades; like Gruendl et al.[36] or Pallet et al.[37] who came up with "beauty ratios". Others, for example, have attempted to develop automated methods to infer aesthetic quality of images (see computational aesthetics field [38]). Typically, these approaches follow a machine learning approach, where large numbers of manually rated photographs are used to "teach" a computer about what visual properties are of relevance to aesthetic quality.

1.5 Goals

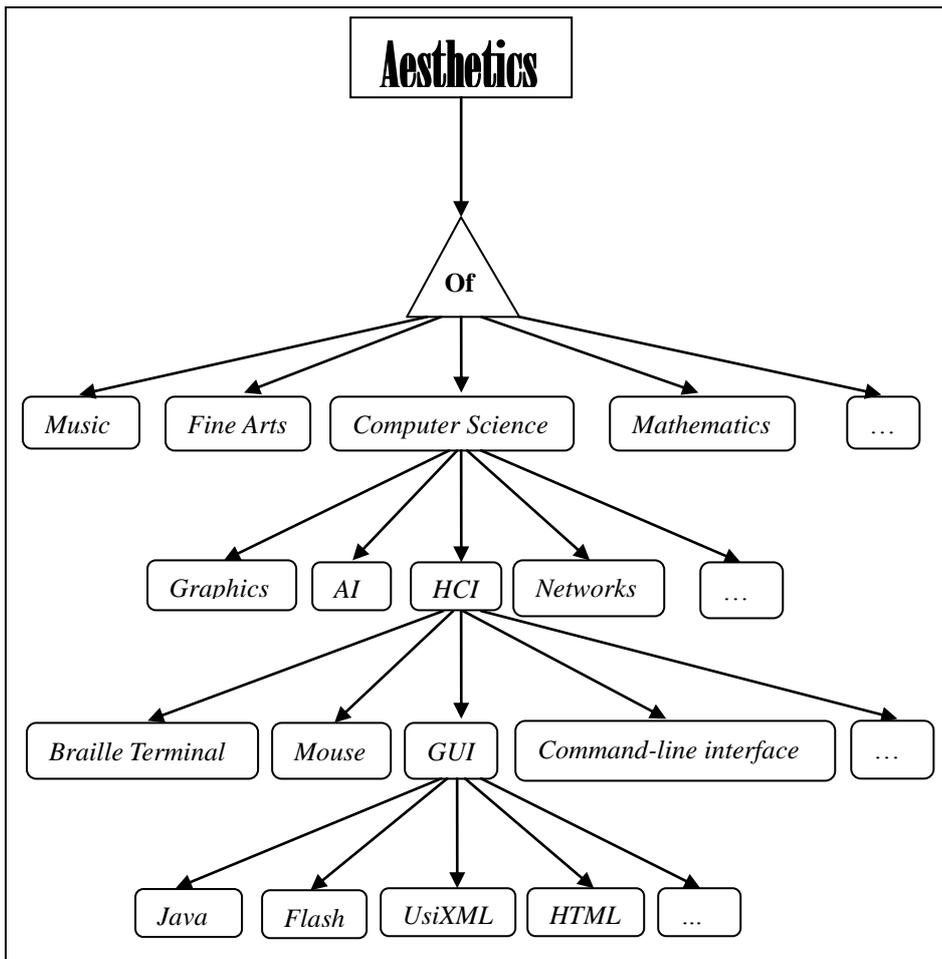


Diagram 1: Application areas of Aesthetics.

In this thesis we will attempt to measure aesthetics. The branch of philosophy of aesthetics can be applied to diverse disciplines, each of which may have their own rules of aesthetics. (*For instance, the aesthetics of cinematography are closely related to still photography, but the movement of the subject(s), or the camera and the intensities, colors, and placement of the lighting are highly important. Sound recording, editing, and mixing are other, highly important areas of film, often closely related with the musical score.*) Unlike Birkhoff who tried to define a universal/interdisciplinary measure of aesthetics, we will restrict ourselves to one particular application area. That is, the field of computer science known as “Human Computer Interaction”. Within this field, we will restrict our experiment to Graphical User Interfaces (GUI) defined in UsiXML (an XML-based markup language for defining user interfaces). See Diagram 1. We believe this is a wise choice because there isn’t much documentation on interdisciplinary approaches and starting our own research from scratch would be unreasonable.

Then, as illustrated in Diagram 2, in our attempt to measure aesthetics we will adopt a philosophical position of Objectivism. This means that we assume that aesthetic properties and qualities belong to objects independently of the observer. Our choice is based on the following reasons: First, we have blind faith that some properties in objects (such as *symmetry*) are universally considered as aesthetic: we believe that an (overly) asymmetric car, face or animal will be displeasing to anyone in any culture, age or gender. Secondly, adopting a subjectivist (i.e. “beauty is in the eye of the beholder”) point of view would force us to conduct brain studies to try and find out the causes of people’s particular preferences. And finally, relativism would require us to combine sociological studies with a machine learning approach in which we would teach an agent (program) to recognize

the features that are considered as beautiful in a particular culture. Then when “asked” if an artifact is aesthetically pleasing, the agent would give several ratings, one for each distinct culture, group or context. These are unexplored research areas and we cannot start from scratch because we simply do not have the skills, time and resources it would take.

Furthermore, although we agree with Hansson [41] that the aesthetic properties and qualities are related not to the “form” of the object only but also to its “function” (see section 2 of chapter III), leading to *aesthetic duality*, we will adopt the *Independence thesis* and spare ourselves the painful process of including “task analysis” in our formulas in order to determine aesthetic quality. So we will assume that the “form” and “function” of an object are judged independently from one another by the observer. We will leave the latter to usability testing and usability research and focus on the former only.

And the most obvious way of measuring the beauty of a form is through mathematical properties such as *symmetry* or *density*; or through other similar guidelines and principles. These are all referred to as “Metrics” or “Formulas” in this work.

So the goal of this thesis can be summed up as follows: *is it possible to measure the aesthetic quality of a UsiXML Graphical User Interface using metrics?*

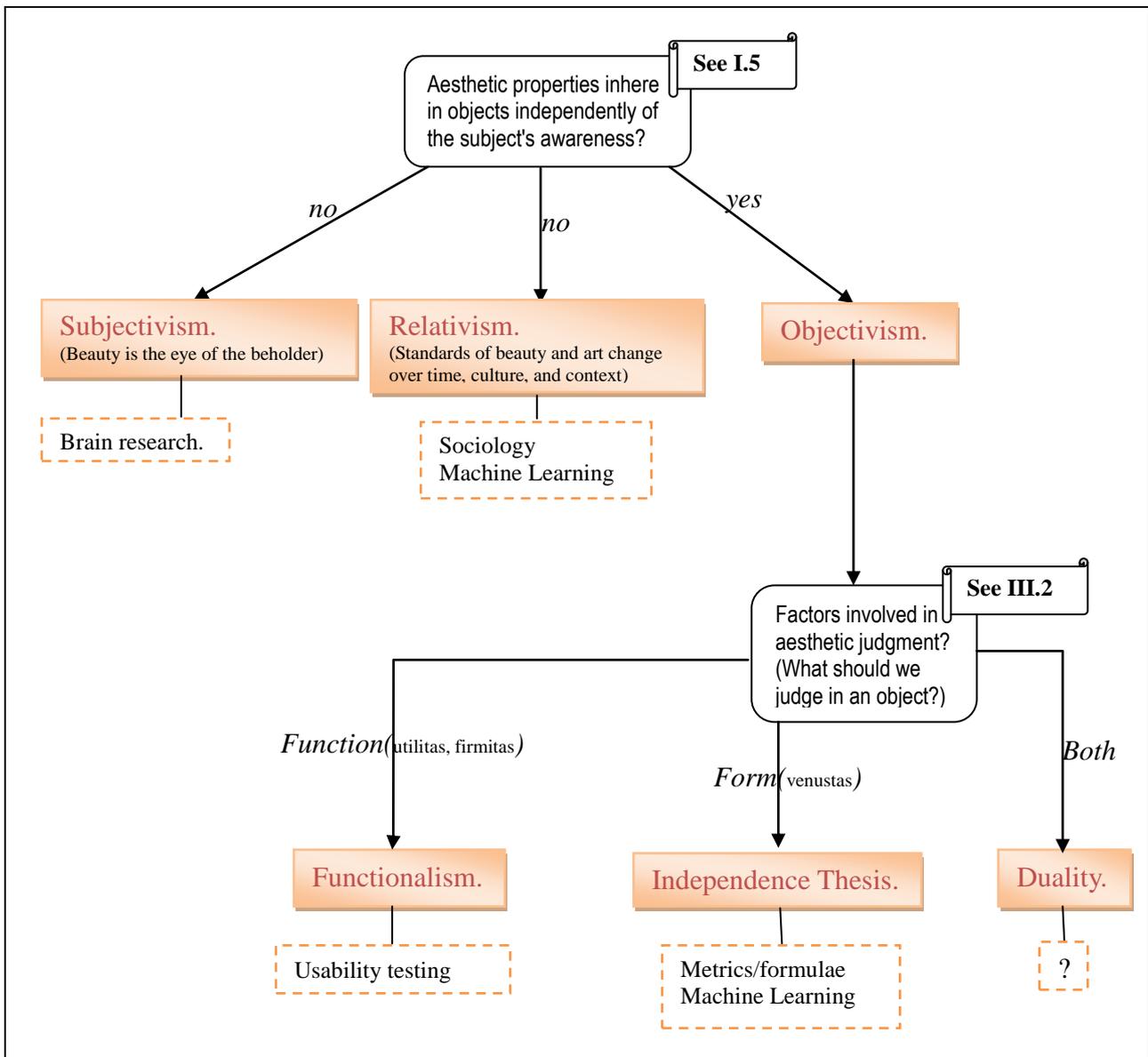


Diagram 2: Aesthetic Judgment; schools of thought and available tools.

Disclaimer: Our ambition is not to follow the *holy grail* of universally measuring aesthetics but merely to provide a tool for designers and usability testers. The results of the automatic evaluation should still be viewed more as guidelines than actual universal measurements. We believe it is certainly true that “no one knows how to measure aesthetic value” [5]. At least not yet.

1.6 Reading plan

In chapter II, we will describe the state of the art in automated evaluation of the aesthetics of Graphical User Interfaces.

Then, chapter III is devoted to the description of our methodology, a study of aesthetic judgment, another study of the origins of the metrics and the process by which we selected the metrics to be implemented.

In Chapter IV, we will describe the implementation of the problem: we’ll first give details about the technologies that we used, and then we’ll describe the architecture of the program.

We’ll describe our tests and report our findings in chapter V. And finally we will discuss the results and draw our conclusions.

Chapter II. State of the art

As mentioned earlier there are several programs at this time that evaluate the aesthetic quality of an interface. Here is a non-exhaustive list thereof:

- Ngo developed as part of a Ph.D. dissertation a software system called “Visit” in which the aesthetic measures described in [6] were installed as part of a tool used to evaluate screen layouts. Visit (which stands for Visitor Information System Implementation Tool) is an experimental design tool for creating “hypermedia systems”.

Visit provides a set of tools that simplify screen design and implementation. It offers the user a library of on-screen objects and supports direct-manipulation interaction in the work area. In addition, Visit has a critic system that helps users design screens. It offers criticism, suggestions, and explanations to help users improve their designs through the use of aesthetic measures.

This work is very similar to ours but its major shortcoming, according to us, is the set of metrics that were selected; not only is that set too small (four metrics is not enough to capture the essence of aesthetics) but also two of the metrics can be done without: *equilibrium* is redundant with respect to balance and *sequence* seems irrelevant to us as aesthetic criteria.

- A program for measuring the fourteen “screen format characteristics” (aesthetic measures) described in [5] was written in the C programming language by Ngo and Byrne. The input to the program is a model example of the screen to be analyzed, drawn over the original screen using a screen editor.

This work is very similar to ours but it has two major shortcomings according to us: on the one hand, forcing the user to redraw a model of the screen is very inconvenient and on the other hand, even though the set of metrics that were selected seems right, we think it is appropriate to question it and come up with our own set of metrics. At first glance, some of the metrics can be done without: *equilibrium* is redundant with respect to balance and *sequence* seems irrelevant to us as aesthetic criteria.

Exemplary ‘good’ and ‘bad’ screens are shown below, along with the corresponding outputs from the program (see fig.4, 5, 6 and 7). The output contains results of the fourteen measurements (balance, equilibrium, symmetry, sequence, cohesion, unity, proportion, simplicity, density, regularity, economy, homogeneity, rhythm, and order and complexity).

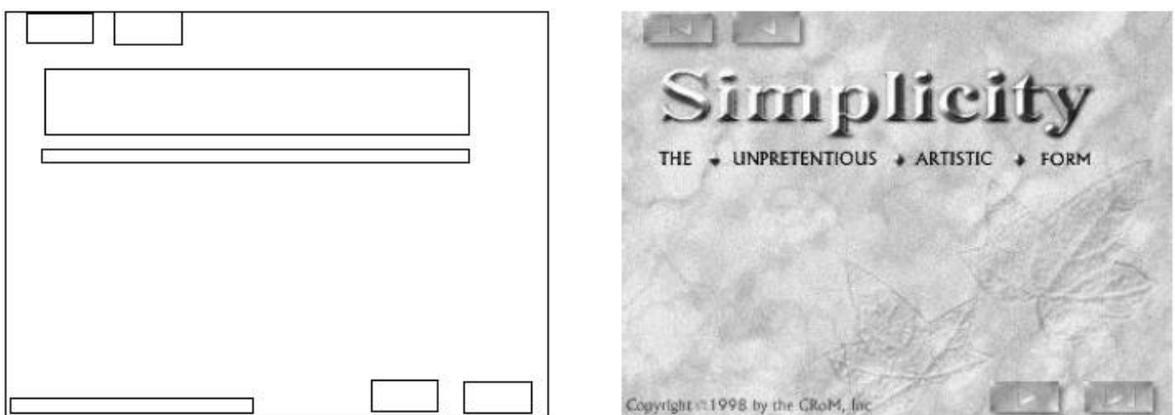


Figure 4: Example of a 'bad' screen.

Measures	Values	Comments
M1 (BM)	0.35711	Unbalanced
M2 (EM)	0.80272	Stable
M3 (SYM)	0.45146	Asymmetrical
M4 (SQM)	0.50000	Random
M5 (CM)	0.67934	Cohesive
M6 (UM)	0.10784	Fragmented
M7 (PM)	0.73442	Proportionate
M8 (SMM)	0.14286	Complex
M9 (DM)	0.41532	Cramped
M10 (RM)	0.08333	Irregular
M11 (ECM)	0.14286	Intricate
M12 (HM)	0.00001	Uneven
M14 (RHM)	0.45306	Disorganised
OM	0.37464	Bad

Figure 5: Output from the program for a 'bad' screen.

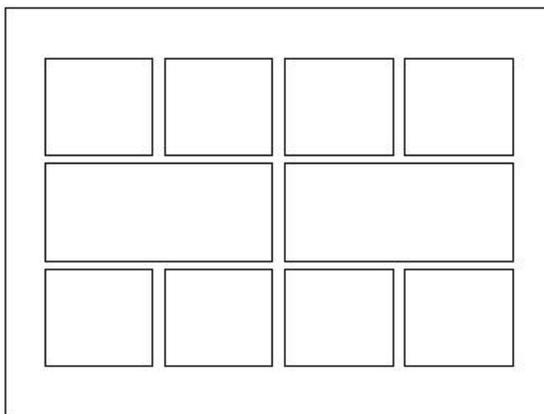


Figure 6: Example of a 'good' screen.

— M1 —		
BMv:	0.634827	Low/Medium (vertically unbalanced: more weight on the left side)
BMh:	0.650946	Low/Medium (horizontally unbalanced: more weight on the top half)
BM:	0.357114	Low/Medium (unbalanced)
— M2 —		
EMx:	-0.10091	High (equal and frame x-coordinates)
EMy:	-0.29364	High (equal and frame y-coordinates)
EM:	0.802724	High (stable)
— M3 —		
SYMv:	0.645608	Low/Medium (vertically asymmetrical)
SYMh:	0.627489	Low/Medium (horizontally asymmetrical)
SYMr:	0.372511	Medium/High (radially symmetrical)
SYM:	0.451464	Medium (asymmetrical)
— M4 —		
SQM:	0.5	Medium (random)
— M5 —		
CMlo:	0.388273	Low/Medium (unequal object and layout aspect ratios)
CMfl:	0.970402	High (equal layout and frame aspect ratios)
CM:	0.679337	Medium/High (cohesive)
— M6 —		
UMform:	0.142857	Low (many object sizes)
UMspace:	0.072816	Low (less space left at the margins)
UM:	0.107837	Low (fragmented)
— M7 —		
PMo:	0.600173	Medium (aesthetically pleasing objects)
PMl:	0.868676	High (aesthetically pleasing layout)
PM:	0.734424	High (proportionate)
— M8 —		
SMM:	0.142857	Low (complex)
— M9 —		
DM:	0.415319	Medium (cramped)
— M10 —		
RMa:	0	Low (many alignment points)
RM _s :	0.166667	Low (many distinct distances between alignment points)
RM:	0.083333	Low (irregular)
— M11 —		
ECM:	0.142857	Low (intricate)
— M12 —		
W:	1.10E+18	
W _{max} :	1.93E+23	
HM:	5.68E-06	Low (uneven)
— M14 —		
RHMx:	0.585517	Medium (rhythmic horizontal arrangement variation)
RHMy:	0.553125	Medium (rhythmic vertical arrangement variation)
RHMa:	0.502181	Medium (disorganised dimension variation)
RHM:	0.453059	Medium (disorganised)

Figure 5(continued)

Measures	Values	Comments
M1 (BM)	0.9963	Balanced
M2 (EM)	1.0000	Stable
M3 (SYM)	0.9985	Symmetrical
M4 (SQM)	1.0000	Sequential
M5 (CM)	0.8012	Cohesive
M6 (UM)	0.8767	Unified
M7 (PM)	0.8686	Proportionate
M8 (SMM)	0.1765	Complex
M9 (DM)	0.8219	Spacious
M10 (RM)	0.7972	Regular
M11 (ECM)	0.5000	Intricate
M12 (HM)	1.0000	Homogeneous
M14 (RHM)	0.9984	Rhythmic
OM	0.8335	Good
— M1 —		
BMv:	0	High (vertically balanced)
BMh:	0.007494	High (horizontally balanced)
BM:	0.996253	High (balanced)
— M2 —		
EMx:	0	High (equal layout and frame x-coordinates)
EMy:	0	High (equal layout and frame y-coordinates)
EM:	1	High (stable)
— M3 —		
SYMv:	0	High (vertically symmetrical)
SYMh:	0.002251	High (horizontally symmetrical)
SYMr:	0.002251	High (radially symmetrical)
SYM:	0.998499	Medium (symmetrical)
— M4 —		
SQM:	1	High (sequential)

Figure 7: Output from the program for a 'good' screen.

— M5 —		
CMlo:	0.716684	High (equal object and layout aspect ratios)
CMfl:	0.885645	High (equal layout and frame aspect ratios)
CM:	0.801165	High (cohesive)
— M6 —		
UMform:	0.9	High (few object sizes)
UMspace:	0.853359	High (less space between objects)
UM:	0.876679	High (unfied)
— M7 —		
PMo:	0.823134	High (aesthetically pleasing objects)
PMl:	0.91404	High (aesthetically pleasing layout)
PM:	0.868587	High (proportionate)
— M8 —		
SMM:	0.176471	Low (complex)
— M9 —		
DM:	0.821875	High (spacious)
— M10 —		
RMa:	0.65	Medium/High (few alignment points)
RM:	0.944444	High (few distinct distances between alignment points)
RM:	0.797222	High (regular)
— M11 —		
ECM:	0.5	Medium (intricate)
— M12 —		
W:	2.31E+12	
Wmax:	2.31E+12	
HM:	1	High (homogeneous)
— M14 —		
RHMx:	0	High (rhythmic horizontal arrangement variation)
RHMy:	0.004796	High (rhythmic vertical arrangement variation)
RHMa:	0	High (rhythmic dimension variation)
RHM:	0.998401	High (rhythmic)

Figure 7(continued)

- In the mid-1990s, Mahajan and Shneiderman [3][4][15] developed a family of tools to help usability testers to efficiently evaluate the consistency of an interface. Their focus was on developing a family of evaluation tools in order to make the evaluation process less cumbersome. They claimed that “inconsistencies in spatial and textual style of an interface designed by several designers may result in a chaotic layout”...that those inconsistencies may have a subtle and negative impact on usability.

So their focus wasn't on aesthetics but rather on usability in itself. Even though some of the metrics they used can in fact be considered metrics of aesthetics, their tool was intended to report ergonomic inconsistencies across dialog boxes of an application whereas we are focusing on evaluating aesthetics; this on one single screen. However, we believe that consistency within one screen is certainly a factor that can influence aesthetic judgment.

- Almost all illustrations in scientific or technical documents employ a large number of textual annotations in the form of labels. In 2005, Hartman et al. proposed several metrics of aesthetic attributes which are used to improve the visual balance of an initial label layout through an efficient real-time optimization process. They implemented several layout styles for labels and integrated real-time label layout algorithms into an

interactive 3D browser. Though users can still specify a preferred layout style for “internal” and “external” labels, the classification and the layout algorithm balance all constraints in order to achieve a readable, unambiguous, aesthetic, and coherent layout.

Their tool includes “functional” (i.e. ergonomic) metrics as well as aesthetic metrics; making their approach close to the notion of *aesthetic duality* that we define in section 2 of chapter III which is the best way to measure aesthetics in our view. However, the area of application of their tool is different from ours.

- In 1987, Perlman [14] proposed an “axiomatic model” of information layout for alphanumeric displays. This model included “a set of axioms (or rules) relating information structure with display attributes”. He also proposed a prototype software system that allowed interactive design and evaluation of screen layouts.

Put simply, the information (labels, fields, relative positions of data, etc...) structure is first represented in a propositional logic-style language and then rules or axioms are applied to the “propositions” to generate a display. As interesting as this approach may be, it seems inappropriate for Graphical User Interfaces because they contain much more information than alphanumeric displays and the ways to (aesthetically) display that information are numerous.

- In 2009, Mirdehghani and Monadjemi [40] developed an automatic system for web pages aesthetic evaluation. Based on the image processing techniques and artificial neural networks, the proposed method was intended to be able to categorize the input web page according to its visual appearance and aesthetic quality. The employed features were multiscale/multidirectional textural and perceptual color properties of the web pages, fed to a perceptron which had been trained as the evaluator. The method was then tested using university web sites and the results suggested that it would perform well in the web page aesthetic evaluation tasks with around 90% correct categorization.

Put simply, they firstly develop an online questionnaire to collect people’s perception and opinions about a known set of university web pages. Then they selected some features which are supposed to be essential in designing beautiful web pages and tried to extract them from their web pages dataset using image processing techniques. When the feature vectors got ready, they fed them to an Artificial Neural Networks (ANN) with appropriate structure as the evaluator, and see the ANN’s opinion about the web page’s aesthetic.

This is a very interesting alternative method! Maybe their only methodological flaw is the use of a homogeneous population (i.e. students). In order to pursue universality the population should be as heterogeneous as possible.

- In 2008, Zain et al. [39] developed and designed Mandarin learning web pages according to the aesthetics values provided by their “Self-Developed Aesthetics Measurement Application” (SDA). This SDA was developed by using Matlab software based on six metrics derived from the model of Ngo et al. [2][5][6] (namely, balance, equilibrium, symmetry, sequence, rhythm, as well as order and complexity). It was developed by using image-processing method, MATLAB Graphical User Interface development environment (GUIDE) and other necessary functions equipped in Matlab software.

This work is very similar to ours but its major shortcoming, according to us, is the set of metrics that were selected; not only is that set too small (five metrics is not enough to capture the essence of aesthetics) but also some of the metrics can be done without: *equilibrium* is redundant with respect to balance, *rhythm* and *sequence* seem irrelevant to us as aesthetic criteria.

Chapter III. Aesthetic Judgment and Study of the Metrics

III.1 Methodology

So we want to measure aesthetic quality, capture the essence of aesthetics. How can that be done? One can either start from scratch or use someone else's findings. Obviously starting from scratch would require much more time, resources and skills than what we have. We cannot, for instance, conduct evolution studies and brain research to try and understand the nature of aesthetic judgments as some have done [27]. The "next best thing" is to trust other people's findings. We don't need to reinvent the wheel!

As we said before aesthetics is originally the branch of philosophy that deals with natural and artistic beauty. So before going any further, we need to scour the available literature, find out what are some of the opinions on the subject and pick one of them. That's what we will do in the next section. We want to determine whether aesthetic quality can be measured. And, if it can, what features should we consider when judging the aesthetic quality of a given object (or artifact). (Needless to say that since philosophy is not an exact science; we can only make a hypothesis and try to prove or disprove it through tests).

In section 5 of chapter I we adopted the position that aesthetics can be measure and in section 2 of this chapter we come to the conclusion that one of the best ways to do this is through metrics/formulas.

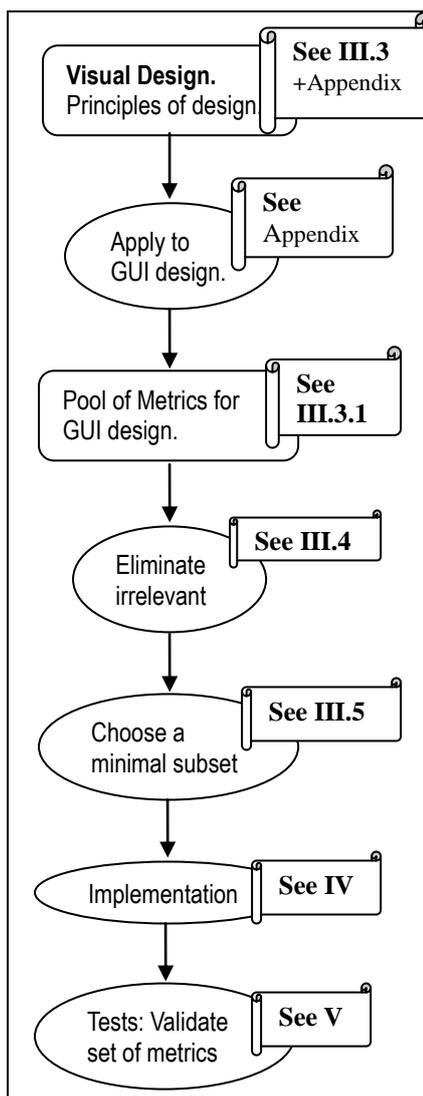


Diagram 3: Methodology.

Several metrics have been proposed in the literature; those metrics rely heavily on techniques and guidelines coming from visual design (indeed, graphic design experts have, through perceptual research, derived a number of principles for defining what comprises a visually pleasing appearance). We are going to choose among them a minimal subset that captures the essence of aesthetics. It goes without saying that choosing the appropriate metrics set is the most important factor in the design of our tool. But how do we choose them? I.e. what method do we use to select the "right" subset of metrics?

There are two possible (empirical) ways of selecting the metrics: we can either choose those that seem right according to the author's subjective and personal way of judging aesthetics, implement them in a tool and then test their validity by conducting an experiment and a statistical study once the tool has been developed (Diagram 3), or we can first conduct an empirical study of the metrics where a heterogeneous population of users would be presented a few GUI designs and then interviewed on the metrics in order to extract a minimal subset that captures the essence of aesthetics. We would then implement them in a tool and conduct no further testing.(Diagram 4)

In both cases we are pursuing universality; in the first case we make the hypothesis that the set of metrics chosen by the author universally captures the essence of aesthetics and then we try to prove or refute it through tests. And in the second case we get a heterogeneous population to choose the minimal set of metrics.

Assuming that the more heterogeneous the population, the more universal the results.

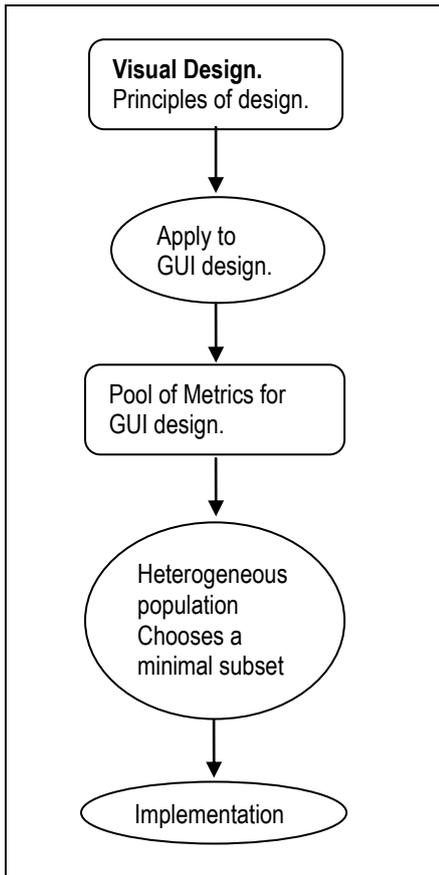


Diagram 4: Alternative methodology.

We chose the first option for the following reason: in our previous work (see Appendix B), users showed impatience during the experiment because of the large number of metrics and the time that it took to first get them to understand the metrics and then to evaluate them on the GUIs. So we often had to hurry as we neared the end of an interview and some answers were given to us expeditiously more out of weariness and impatience than real thoughtful analysis.

So we believe that in order to arouse people’s interest and get them to passionately participate we need a different, more sensational approach. People will be more excited to test a “tool that measures beauty” than to “help choose formulas that will help build such a tool”. Thus, we believe this approach will eliminate the bias of boredom and tiredness. And we’ll also be able to attract a really heterogeneous population and not just our relatives and friends.

So here is how we will do this (see Diagram 3): In section III.4, we will first discard the metrics that are obviously irrelevant. Then in section III.5, we will study the rest one at a time using various examples of GUIs and decide subjectively which ones are redundant or useless and keep a minimal subset that we will

implement in our tool in chapter IV. And finally, in chapter V, we will conduct tests and an experiment with users to evaluate the effectiveness of our tool and the subset of metrics that we proposed in assessing aesthetics. This will allow us to draw conclusions and discuss future work in the last chapter.

III.2 Aesthetic experience

In chapter I, section 4 and 5, we have already defined the philosophical positions of subjectivism, relativism and objectivism. And we explained why we chose the objectivist point of view. Now (within an objectivist view of the world), we still need to find out what constitutes an aesthetic experience: in any field, what makes an artifact pleasing or displeasing to the subject (or viewer) aesthetically? For this purpose, let’s first make a systematic classification of artifacts based upon Vitruvius’ theory of architecture. (Diagram 2 graphically summarizes this discussion).

III.2.1 Classifying artifacts

“The first known systematic theoretician of architecture, Vitruvius (first century BC), argued that architecture must satisfy three distinct requirements: *firmitas* (strength)—which covers the field of statics, construction, and materials; *utilitas* (utility)—the use of the building and its functioning; and *venustas* (beauty)—the aesthetic requirements. Although architectural theories have since evolved considerably, Vitruvian principles still hold much intuitive and theoretical appeal to this date”. [10]

There are many definitions of the concept of “Form”, but for our purposes here let us define it as simply the “shape” of an artifact. So the artifact itself is the form and we will use both terms

interchangeably. And let us define “Function” as the *utilitas* of the artifact. However, considering the fact that if something has no *firmitas* at all it cannot be useful, our definition of “Function” will also include *firmitas*.

Now we can reasonably classify the set of all possible forms based on their intended purpose and according to two basic attributes: Function and Beauty. So there are three theoretical categories of objects:

1. Functional (*venustas* is unimportant)
2. Hedonistic (*venustas* is their only purpose)
3. Function and Beauty (both function and *venustas* are important)

The first category contains artifacts of any field that are only functional. Here, aesthetic considerations are inexistent and irrelevant or if they do exist then they are a very unimportant factor of design. For example, does it matter to the brick maker whether a brick in itself is beautiful? We expect artistic photography to be beautiful but what about functional photography? Does the *venustas* of an ID photo really matter? (That is, the photo in itself. Not the looks of the person).

Note that it has been argued that there is beauty even in the fields that seem austere at first sight. For instance, Bertrand Russell expressed his sense of “mathematical beauty” in these words: Mathematics, rightly viewed, possesses not only truth, but supreme beauty — a beauty cold and austere, like that of sculpture, without appeal to any part of our weaker nature, without the gorgeous trappings of painting or music, yet sublimely pure, and capable of a stern perfection such as only the greatest art can show. The true spirit of delight, the exaltation, the sense of being more than Man, which is the touchstone of the highest excellence, is to be found in mathematics as surely as poetry. [42]

Of course, we agree that anything can be perceived as “beautiful”. But we must distinguish between pure artistic beauty (*venustas*) and other forms of beauty. Remember that our classification is based on the purpose of an artifact. Artistic beauty is beauty for itself, beauty as a purpose. Mathematical proofs and algorithms might be perceived as beautiful but being beautiful is not their only purpose, they are also expected to solve problems. An abstract sculpture or painting seeks beauty for beauty; it has no other purpose.

The second category contains artifacts whose sole purpose is to be pleasing to the senses. To look, smell, sound and (possibly) feel beautiful. Objects like abstract sculptures and paintings or ornamental gardens don’t seek to satisfy the *utilitas* requirement as much as the *venustas*.

Arguably this kind of artifact does actually have a function: to please the senses. So by being beautiful, they are actually fulfilling their function.

But once again let us make a distinction here, between practical function (*utilitas* and *firmitas*) and aesthetic function (*venustas*).

The third category contains forms that are created with the intention to satisfy primarily the criterion of practical function, but whose *venustas* proves to be an important factor, a determinant of user satisfaction, among other things. It’s easy to find examples of such artifacts. Consider cars, for example, or clothes, architectural works, etc.

III.2.2 Aesthetic Judgment

Now that we have classified forms based on their intended purpose, let’s try to answer our initial question: what makes an artifact aesthetically pleasing? Or more precisely, what criteria do we consider in our aesthetic appraisal of a form? We already know that “what is beautiful is usable”. What about the opposite? Is what is usable perceived as beautiful? Is our aesthetic judgment only

influenced by *venustas* or do *utilitas* and *firmitas* influence our perception as well?

If the object belongs to the first category, the question is pointless since these kinds of objects are made without any aesthetic consideration whatsoever. Artistic beauty is not part of their intended purpose nor is it part of their use. A brick layer, for instance, is not interested in the aesthetic quality of a brick. However, like we said before, if the artifact is considered in a different context of use. If, for example, the brick layer takes the brick to a museum, then it might qualify for aesthetic appraisal. But, otherwise, the question is pointless in this category.

If the object belongs to the second category, then the question is of paramount importance; since the designers, artists and critics are all interested in knowing what makes a painting, for example, good or bad, or a song, etc. And since the form is shaped according to the *venustas* requirement alone, the question boils down to: what constitutes good *venustas*?

But if the artifact belongs to the third category, then the form is no longer shaped based on *venustas* requirements alone but also on function (*utilitas* and *firmitas*). This combination gives rise to the crucial issue of how the two types of value relate to each other. In our aesthetic appraisal of such a form do we consider only its *venustas*? Does function influence our aesthetic judgment as well? Is what is usable beautiful? Is practical function independent of aesthetics or are they in some way connected?

Hansson [41] identifies three currents of thoughts: *aesthetic functionalism*, *independence thesis* and *aesthetic dualism*.

The “**functionalist**” viewpoint claims that the aesthetic properties of an object [of the third category] depend on its functionality. This was a central issue in the *functionalist movement* in architecture and design, one of the most influential artistic and cultural movements of the early twentieth century. Their most famous slogan was "form follows function," whose originator was the American architect Louis Sullivan. When introducing it, he made an analogy with the forms and functions one can find in nature. The form of the eagle's wing has been determined by its function, and the same applies to other objects in nature. "Form ever follows function. This is the law. Shall we, then, daily violate this law in our art?"

The most far-reaching variant of functionalism can be called the "reduction thesis." It consists in the claim that there is nothing to aesthetics (beauty) in addition to what follows from function. This has also been called "austere functionalism;" it implies that aesthetic considerations are altogether unnecessary; since aesthetics will be automatically taken care of if function is dealt with adequately. This view was expressed by Wilhelm Wagenfeld in an article published in 1937. He described how he had developed a new lemon squeezer by carefully trying out different forms in his "artistic laboratory" until the most functional construction was found. This construction was then selected for mass production. The experiments performed on lemon squeezers were in his view an aesthetic activity. "Being useful also means being beautiful." [41]

So according to the “functionalist”, in our aesthetic judgment of an artifact of the third category, we need only consider the *utilitas* of the form. This implies that if we are to build a tool that evaluates aesthetic quality we need only ask ourselves: is it useful? If it is, then it’s beautiful!

Hansson [41] argues that the reduction thesis may be appealingly simple, but it is fraught with difficulties. To begin with, it has problems in dealing with pure art. Consider two objects that are both made by the same glass-blower. One is a vase that is used to keep flowers in, and the other an artwork called "pillar" that cannot be used as a vase since it has a hole in its bottom. It so happens that the two objects are strikingly similar to each other. According to the reduction thesis, the aesthetic properties of the vase are entirely determined by its function as a vase. The "pillar," however, has no function (or rather, no function that is prior to and independent of its aesthetic properties). Therefore its aesthetic properties cannot be derived in the same way as those of the

vase. According to ordinary aesthetic intuitions, there is much in common between the aesthetic criteria that we apply to the vase and those that we apply to the "pillar." A line that we find graceful or elegant in one of them would probably be found to have the same property in the other. The reduction thesis makes it difficult to account for this, since it subsumes the aesthetic properties of the vase under its function that is not at all shared by the "pillar". Another problem for the reduction thesis is that practical function underdetermines form or, in other words, that objects that differ widely in their physical appearance can realize one and the same practical function with equal efficiency. Two soup-plates may be equally useful (functional) but yet look very different. According to the reduction thesis, they must then satisfy aesthetic demands to the same degree, which is implausible to say the least.

So functionalism and its "reduction thesis" variant seem utterly incorrect to us.

Hansson [41] says the **independence thesis** claims the opposite. According to this standpoint, aesthetic and practical values are completely independent of each other. In the early twentieth century, the independence thesis was defended by Hermann Muthesius in his attacks on the functionalist movement. He said: "It is a mistaken idea to believe that it is fully sufficient for an engineer to see to it that the house, instrument, or machine that he is constructing satisfies a purpose. It is even more mistaken to believe in the saying, so often heard nowadays, that if it fulfills a purpose then it is also beautiful. Usefulness as such has nothing to do with beauty. Beauty is a matter of form and nothing else, usefulness is purely the matter of being serviceable." It is not difficult to use examples to show how the independence thesis goes too far in the direction opposite to that of the reduction thesis. Perhaps the clearest counterexamples are those that refer to the beauty of abstract objects, such as mathematical proofs. A mathematician who called a proof beautiful would almost certainly retract or at least moderate that statement if the proof turned out to contain irreparable mistakes. Similarly, someone who admires the beauty of a scientific experiment does this under the presumption that it actually works. When a computer program is called beautiful, this is usually because of its power to accomplish a lot with surprisingly small means. This type of beauty was called "machine beauty" by David Gelernter.

Examples can also easily be found that refer to concrete, technological objects. Few of us would admire the beauty of a bicycle with oval wheels or a chair that falls apart if one sits on it. Or rather, if we appreciate them aesthetically, then we appreciate them as non-functional works of art, not as a bicycle or chair. When we judge a chair aesthetically, we typically make assessments of it that relate to its function as a chair. That an object is, for instance, a "beautiful chair" does not only mean that it is both beautiful and a chair. It means that it is beautiful as a chair.

So the independence thesis can be discarded as well.

Aesthetic duality is an intermediate standpoint between the reduction and the independence theses: When an object has a purpose or practical function, then some but not necessarily all the aesthetic judgments that can legitimately be made about the object refer to that function. So, practical function influences our aesthetic judgment. Take the resumes in figure 8, 9 and 10, for example. They all have an excellent *venustas* but they also have a very poor *utilitas*. They are beautiful. But they are not beautiful *as* resumes. A more usable (or user friendly, or ergonomic) form like in figure 11 appears to be much more aesthetic for a CV.

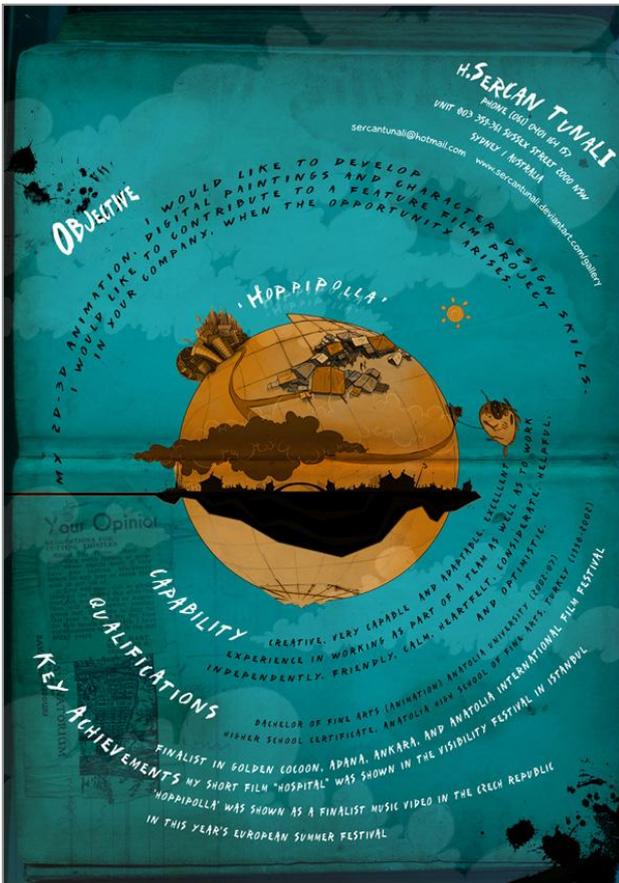


Figure 8: A graphic designer's C.V.



Figure 9: A graphic designer's C.V.

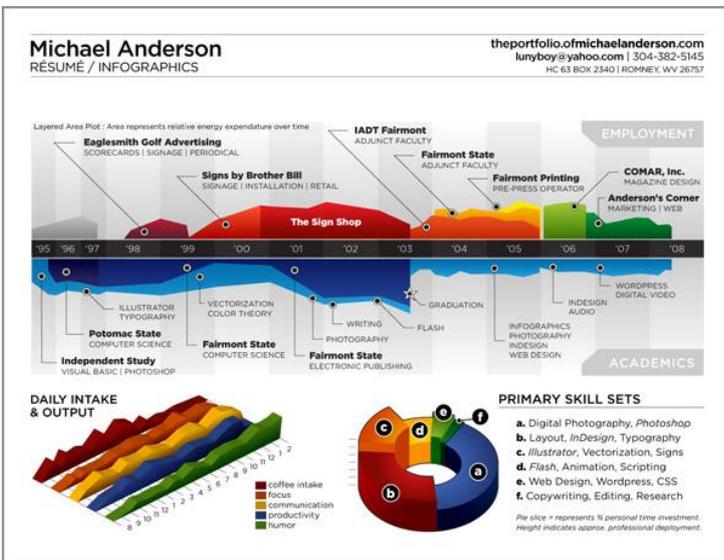


Figure 10: A graphic designer's C.V.

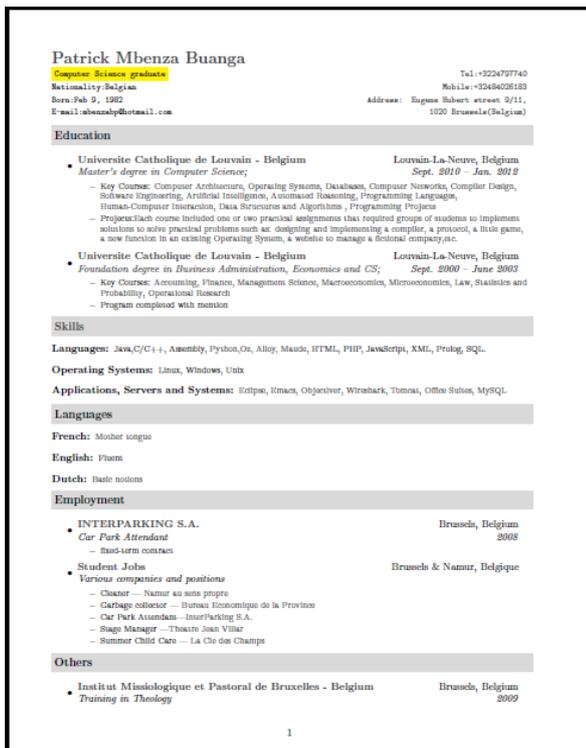


Figure 11: A more classic C.V.

What point did we try to make in this section? We demonstrated in this section that our aesthetic judgment of a form of the third category (e.g. a Graphical User Interface) is inevitably influenced by both its *venustas* and its function. This means that if we want to develop a tool that evaluates aesthetic quality, we have to take into account those two aspects. We cannot consider them separately.

III.2.3 The GUI artifact

Consider, for our purpose, the field of information systems. In chapter one, we have established the importance of aesthetics in HCI. There is little doubt that Graphical User interfaces are artifacts of the third category according to our classification in section one. So they must satisfy all three criteria *firmitas*, *utilitas* and *venustas*.

Much of the work in the field of IT relates to the soundness and the robustness of the artifacts created by professionals in the field (*firmitas*). Traditionally, the various computing and information technology disciplines have emphasized areas related to the firmness, correctness, stability, and internal logic of their products. [10]

The second requirement, *utilitas*, is addressed by a certain stream in MIS research and by a large segment of the human-computer interaction (HCI) community. It deals with the ways in which information technology can be designed to meet individual and organizational needs with regard to the systems' functionality and ease of use.

Until very recently, however, the third Vitruvian requirement, *venustas*, was almost completely absent from research in the various IT disciplines.

Lee [18] says: "there is an unspoken battle currently raging among those who develop user interfaces. It's the battle between usability experts and graphic designers. On one hand, usability experts believe a user interface should be simple and that tasks should be able to be completed quickly and effectively. On the other hand is the graphic designer who believes the user interface

should be visually interesting, aesthetically appealing, unique and also entertaining. The web is just too big for either one of these disciplines to prevail. Some sites need to be artistic and entertaining, while others need to be made as simple and intuitive as possible. The future of user interface design is to create sites that are both easy to use, as well as being visually appealing”.

Now, we saw in section 2 that *aesthetic duality* was the best way of describing the aesthetic judgment of third category artifacts. This means that a tool that aims at imitating the human aesthetic judgment of GUIs shouldn't only implement the universal rules of *venustas* but also take into account the usability of the GUI (this process is called “task analysis”) and the way both aspects are related! This is not a trivial task.

Some of the metrics proposed by the authors do in fact include task analysis. But they never explain how that analysis should be made. For instance, Vanderdonckt and Gillo [16] define *economy* as “the frugal and judicious use of widgets” and “no widget that is extraneous to the user's task appears in the layout”. But how can one know for sure what the user's task is, on the exclusive basis of the look of the interface?

So, even though aesthetic duality is the most accurate approach according to us, we are going to adopt an *independent thesis* position. Just so we can avoid the daunting task of including user's task analysis: so we will assume that only the *form* of a GUI is considered during an aesthetic experience.

In that case our initial question becomes: What is the most accurate way of automatically assessing the *venustas* (form) of a GUI? From our research we could find only two methods: metrics and machine learning. Obviously every form in nature can be rated objectively in terms of *symmetry* or *asymmetry*, for example. A few authors provide us with other such features (metrics) that can be objectively measured on objects and that contribute to their beauty.

Another approach is to collect the ratings of a (very) large and heterogeneous crowd and then teach an Artificial Neural Network how to rate screens based on the data collected. [40] Both methods are equally accurate in our opinion. Here is why: no observer can deny the law of gravity described by physics. It transcends particular opinions because it can be demonstrated objectively. But the laws of aesthetics (even the most basic ones like *symmetry*) still seem to need to get the observer's agreement: it seems like there is no way to prove an observer who prefers *asymmetry* wrong. So in any case, we need a universal and absolutely objective observer to tell us what features make a beautiful form beautiful. The closest instance of such an observer is a crowd (we will hereafter refer to it as the “God-crowd”) so large and heterogeneous that the bias of cultural conditioning and particular whims are removed. In the case of Machine learning, the ANN is taught the laws of aesthetics by the God-crowd. And in the case of metrics, the God-crowd either chooses the metrics to be implemented (second method described in section III.1, see Diagram 4) or validates the metrics that have been chosen by another (first method described in section III.1, see Diagram 3).

Nonetheless metrics are a more straightforward and intuitive method. And they do not require mastering any special field of expertise.

III.3 A Brief Study of the Metrics

III.3.1 Origin of the Metrics

The set of metrics that we are going to consider are those that have been proposed by Ngo et al. [2][5][6], Vanderdonckt and Gillo[16][17], Shneiderman et al. [3][4][15], Perlman [14] and Hartmann [19].

These authors' works rely heavily on techniques coming from visual design. *Visual design* in

general is interested in arranging information items (e.g., text, images, diagrams, pictures, tables) in such a way that it is visually attractive, perceptive and easily understandable. Visual design issues are raised in many domains of human activity such as user interface design, documentation development, presentation design, and graphic layout.

Ngo et al. use mainly Galitz' book on design and layout, *The Essential Guide to User Interface Design* and other similar works. They also use texts that discuss theories of design in both fine art and commercial art.

Shneiderman et al. worked closely with General Electric Information Services and a group of researchers at the University of Maryland to design and develop their metrics.

The researchers at the University of Maryland generated a list of approximately 40 metrics which were constructed after reviewing the relevant literature at the time (namely Tullis(1988), Streveler and Wasserman (1987), Coll and Wingertsman(1990), Kim and Foley(1993), Vanderdonckt and Gillo(1994), Bodart et al.(1994), sears(1993,1994), Combert and Maltby(1995), and Mullet(1995)), consulting with colleagues and using their GUI evaluation experience. A similar effort was taken on the GE side, where they brain-stormed and proposed their metric set based on their commercial software development experience. The two lists had many similar items and the lists were grouped into categories such as spatial layout, alignment, clustering, cluttering, color usage, fonts, attention getting, etc. The metric set was then revised several times after evaluating a series of interfaces. The metrics that were ineffective were removed and others were redefined and new metrics were added.

In the next section, we are going to give more details about visual design. And then in section 3, we will draw a parallel between visual design and GUI design.

III.3.2 Visual design in a nutshell

Visual Design is the organized arrangement of one or more *elements* and *principles* for a purpose. The *elements* and *principles* of design describe fundamental ideas about the practice of good visual design that are assumed to be the basis of all intentional visual design strategies. The *elements* form the “vocabulary” of the design, while the *principles* constitute the broader structural aspects of its composition. Awareness of the *elements* and *principles* in design is the first step in creating successful visual compositions. These *principles*, which may overlap, are used in all visual design fields, including graphic design, industrial design, architecture and fine art.[25][26]

Elements of design

Design elements are the building blocks or basic units in the construction of a visual image. Design elements include:

- ***Line***
It is a mark with greater length than width. Lines can be horizontal, vertical or diagonal, straight or curved, thick or thin.
- ***Shape***
Anything that has height and width has shape. Unusual shapes can be used to attract attention. There are basically three types of shapes. Geometric shapes, such as triangles, squares, rectangles, and circles, are regular and structured. These shapes work very well as building blocks for graphic design. Natural shapes, such as animals, plants, and humans, are irregular and fluid. Abstracted shapes, such as icons, stylized figures, and graphic

illustrations, are simplified versions of natural shapes.

- **Texture**

Texture is the look or feel of a surface. Visual texture creates an illusion of texture on a printed publication or web page. Patterns, such as the images printed on wrapping paper, are a type of visual texture.

- **Space**

Space is the distance or area between or around things. Space separates or unifies, highlights, and gives the eye a visual rest.

- **Size**

Size is how large or small something is. Size is very important in making a layout functional, attractive, and organized. It shows what is most important, attracts attention, and helps to fit the layout together.

- **Color**

Color in layouts can convey moods (fig.12), create images, attract attention, and identify objects. When selecting colors for a publication or a web page, designers should think about what they want the color to do and what is appropriate for their purpose.

Color	Psychological Associations
Red	Passion, strength, bravery, energy, romance
Blue	Loyalty, dependability, confidence, productivity
Green	Nature, fertility, life, renewal, harmony, luck, wealth
Yellow	Joy, happiness, wealth, hope, friendship
Orange	Energy, enthusiasm, flamboyance, playfulness
Purple	Royalty, wisdom, creativity, mystery, ceremony
Brown	Comfort, depth, nature, stability, poverty, dependability
Black	Power, sophistication, formality, elegance, style
Gray	Balance, subtlety, respect, formality
White	Purity, neutrality, simplicity, innocence, cleanliness

Figure 12: The psychology of color.

- **Value**

Value is the lightness or darkness of an area. For instance, in the spectrum from black to white value is the many shades of gray in between. Each shade on this spectrum has a value, from the very lightest to the very darkest. Value separates, suggests mood, adds drama, and creates the illusion of depth.

Principles of design

The principles of design help to determine how to use the design elements, to combine them into a good layout. The principles of design suggest effective and pleasing ways to arrange text, graphics, etc... on a page.

The principles of design are varied. They differ both between the schools of thought that influence design, and between individual practicing designers. Design principles include, for example:



Balance is the distribution of the visual weight of objects, colors, texture, and space. If the design was a scale these elements should be balanced to make a design feel stable. In symmetrical balance, the elements used on one side of the design are similar to those on the other side; in asymmetrical balance, the sides are different but still look balanced. In radial balance, the elements are arranged around a central point and may be similar.



Emphasis is the part of the design that catches the viewer's attention. Usually the artist will make one area stand out by contrasting it with other areas. The area will be different in size, color, texture, shape, etc.



Movement is the path the viewer's eye takes through the artwork, often to focal areas. Such movement can be directed along lines edges, shape and color within the artwork.



Pattern is the repeating of an object or symbol all over the artwork.



Repetition works with pattern to make the artwork seem active. The repetition of elements of design creates unity within the artwork.



Proportion is the feeling of unity created when all parts (sizes, amounts, or number) relate well with each other. When drawing the human figure, proportion can refer to the size of the head compared to the rest of the body.



Rhythm is created when one or more elements of design are used repeatedly to create a feeling of organized movement. Variety is essential to keep rhythm exciting and active, and moving the viewer around the artwork. Rhythm creates a mood like music or dancing.



Variety is the use of several elements of design to hold the viewer's attention and to guide the viewer's eye through the artwork.



Unity is the feeling of harmony between all parts of the artwork creating a sense of completeness.

Typography

Typography is the art and technique of arranging type in order to make language visible. The arrangement of type involves the selection of typefaces, point size, line length, leading (line

spacing), adjusting the spaces between groups of letters (tracking) and adjusting the space between pairs of letters (kerning). Type design is a closely related craft, which some consider distinct and others a part of typography; most typographers do not design typefaces, and some type designers do not consider themselves typographers.

Typography is performed by typesetters, compositors, typographers, graphic designers, art directors, comic book artists, graffiti artists, clerical workers, and anyone else who arranges type for a product. Until the Digital Age, typography was a specialized occupation. Digitization opened up typography to new generations of visual designers and lay users.

III.3.3 GUI design in a nutshell

The metrics proposed come from visual design; they are expanded to user interface design. So let's draw a parallel between the two fields:

Elements of GUI design

Here, the elements consist of interaction objects and interactive objects. Interaction objects (IO), also called widgets or controls, encompass static objects (e.g. labels, separators, group boxes) and dynamic objects (e.g. edit boxes, radio boxes, option boxes). Interactive objects cover every other kind of object that a *multimedia* user-interface could virtually display: static icons, drawings, pictures, images, sketches, video sequences, graphics, etc. Each of these objects allows some special interaction with the user. For instance, an image of the human body may include hot spots for defining different sensitive regions of the body in order to be selected, displayed, explained or zoomed. Some images can be extracted from a video sequence in order to be analyzed. Interaction and interactive objects will be further referred to as IO. [17]

But the elements of visual design (see. previous section) are also elements here because the content of images (e.g. background images), the colors of buttons, frames, backgrounds, texts, the shapes, etc. play an important role.

Principles of GUI design

When the designer sketches the components of a user interface, the first thing to do is to select appropriate interaction and interactive objects according to the user's task. The second activity is to determine the basic layout of these selected objects also according to the user's task. And for this, the designer should use guidelines coming both from ergonomics (*utilitas*) and Aesthetics (*venustas*):

➤ GUI Ergonomics

○ Utility

The GUI effectively and efficiently accomplishes the tasks for which it was designed. It is correct with respect to the functional objectives of the application.

○ Usability

Ease of use, learnability of the GUI. The elegance and clarity with which the interaction with the application is designed. Bastien and Scapin [], for instance, define eight “ergonomic criteria”:

- **Consistency**
Certain aspects of an interface should behave in consistent ways at all times for all screens; terminology, icons, colors... should be consistent between screens or within a screen.
- **Guidance**
refers to the means available to advise, orient, inform, instruct, and guide the users throughout their interactions with a computer (messages, alarms, labels, etc.), including from a lexical point of view.
- **Workload**
concerns all interface elements that play a role in the reduction of the users' perceptual or cognitive load, and in the increase of the dialogue efficiency.
- **Explicit Control**
concerns both the system processing of explicit user actions, and the control users have on the processing of their actions by the system.
- **Adaptability**
refers to its capacity to behave contextually and according to the users' needs and preferences.
- **Error Management**
refers to the means available to prevent or reduce errors and to recover from them when they occur. Errors are defined in this context as invalid data entry, invalid format for data entry, incorrect command syntax, etc.
- **Compatibility**
refers to the match between users' characteristics (memory, perceptions, customs, skills, age, expectations, etc.) and task characteristics on the one hand, and the organization of the output, input, and dialogue for a given application, on the other hand.
- **Significance of Codes**
qualifies the relationship between a term and/or a sign and its reference. Codes and names are significant to the users when there is a strong semantic relationship between such codes and the items or actions they refer to.

➤ GUI Aesthetics

- Aesthetic measures. (e.g. Ngo et al.)
- Visual techniques. (e.g. Vanderdonckt and Gillo)

III.3.4 Limits of the metrics

After having studied the metrics, there are some observations that we can already make:

- The metrics that have been proposed are purely visual, resting on aesthetic and psychological factors and disregarding human aspects deriving from user requirements;
- Effective visual design should rely on task analysis. Our metrics are considered separately from task and user's requirements. This disconnection should not be forgotten. Consider, for example, the contact pages in fig.13 and 14. They use very expressive metaphors that are related to the task and user requirements (good ergonomic *compatibility*);

yet they may receive negative evaluations by the metrics; such as *misalignment* for fig.13.



Figure 13: A website's contact page.



Figure 14: A website's contact page.

- Applying these metrics in a general fashion does not preclude that users have to be avoided. Conversely, specifying layouts should follow one actual trend characterized by the cooperative participation of the users. The aim of the metrics is simply to keep the designer's mind conscious to the visual choices made with these users;
- The general composition of the layout of a screen includes aspects that are left out by the metrics:

- images are considered as an indivisible whole; as a unique element or IO. However, the content of background images can be used to influence the general aesthetic and ergonomic experience. (e.g. Fig .15)

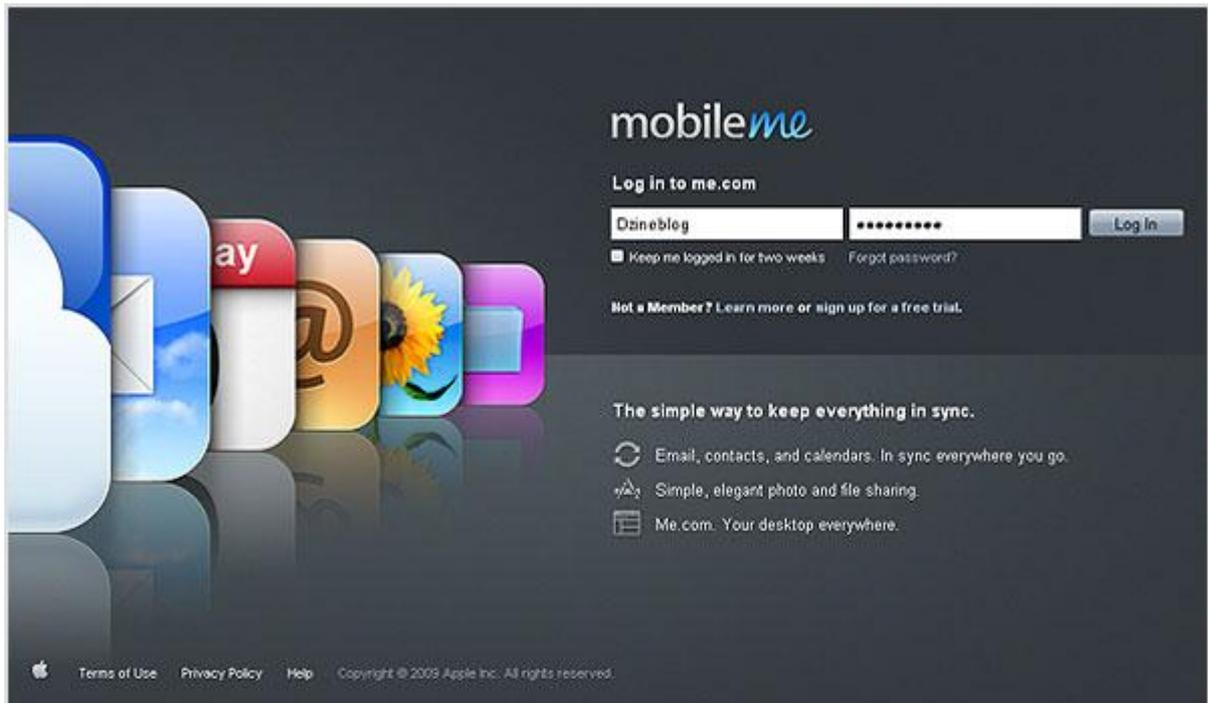


Figure 15: The background image influences the composition of the layout.

- Typical techniques of multimedia applications are not considered: for instance **animation**, tiling, window sequencing, etc.
- the metrics should consider other aspects coming from visual design; such as typography or color psychology (the colors in fig.16 play an important role in the aesthetic experience) or the study of textures and patterns like in dressmaking and fashion design (fig.17) ;

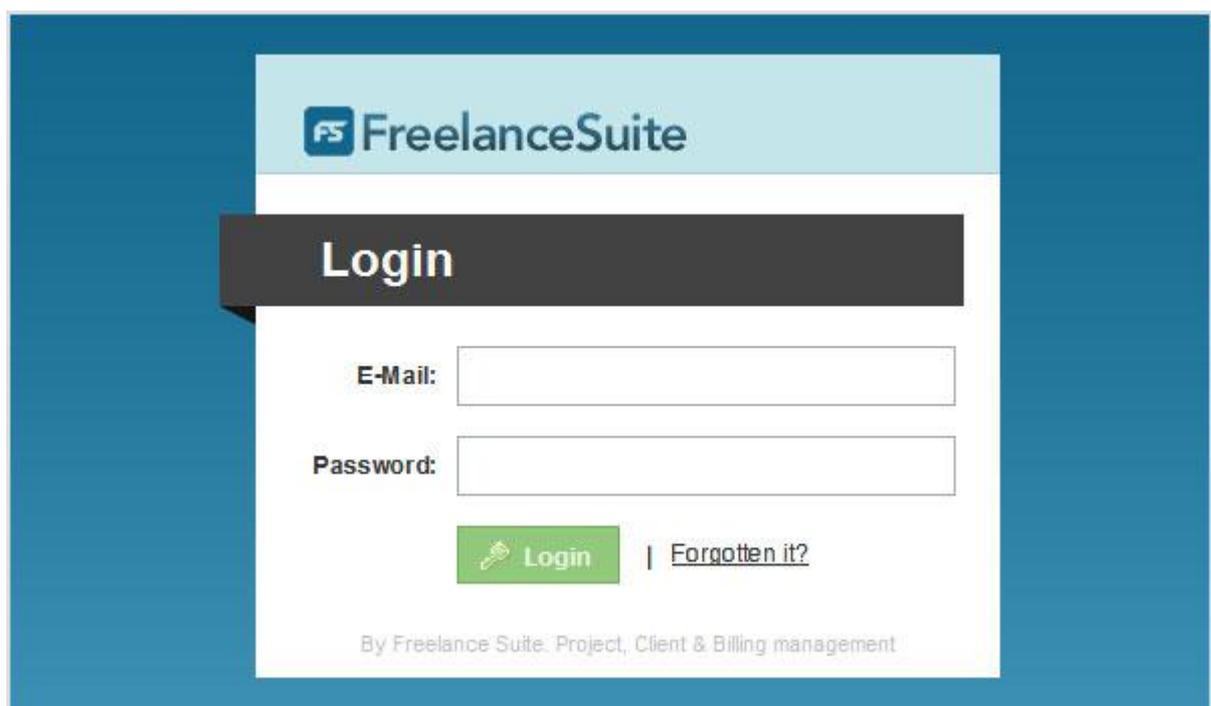


Figure 16: login page with interesting colors.

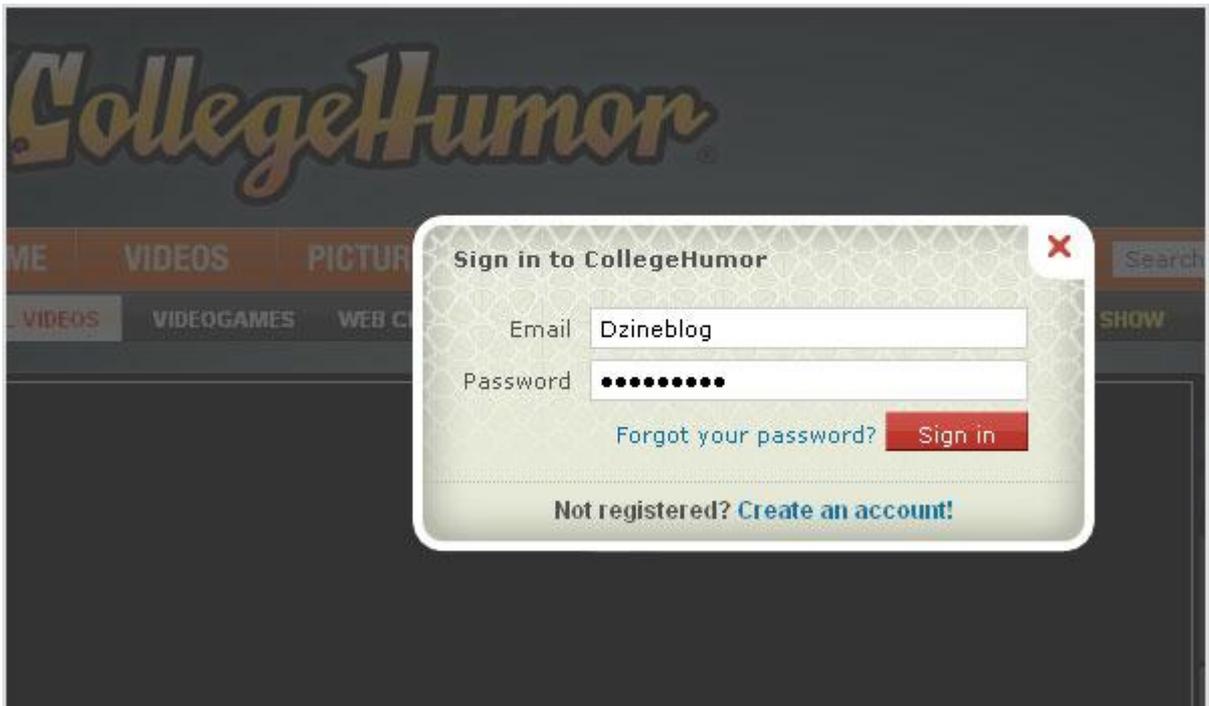


Figure 17: The texture of the dialog box influence our aesthetic judgment.

- The metrics assume that the composition of the screen is two dimensional. This leaves out perspective and 3D as in fig.18.



Figure 18: Contact page with perspective.

III.4 Eliminating irrelevant Metrics

In this section we are going to “prune” our pool of metrics, reduce the size of the set of metrics that will be studied and considered for implementation by eliminating those that are obviously

In graphical user interfaces, the density is calculated by dividing the number of lighted pixels by the total number of available pixels. Streveler & Wasserman also measure the field density which is the total amount of fields (static or dynamic) in the layout, and the box density, which is the total amount of visual groups whether surrounded or not. Tullis recommends that layout density should not exceed 25%. Horton recalls us that density of a well-designed paper page is located around 40%. [16]

- Comments: none
- Importance/Weight: medium

❖ Formula

$$DM = 1 - \frac{\sum_i^n a_i}{a_{frame}} \in [0,1]$$

Where a_i and a_{frame} are the areas of object i and the frame; and n is the number of objects on the frame (or interface).

❖ Motivation

The interface in fig.18 is very unaesthetic (in the authors view). After taking a closer look at it, we figured this was because of three main problems:

- Misalignment of widgets (text fields, drop-down list).
- Inconsistency in the positions and formats of the labels
- A cluttered and overcrowded layout

Fig. 19 has as many widgets as Fig.18 but alignment and consistency make it a beautiful interface to behold. Fig.20 is misaligned and inconsistent but still beautiful to behold (in the authors view). Why? Just because there is a lot of space. It is not crowded. Its density is very low.

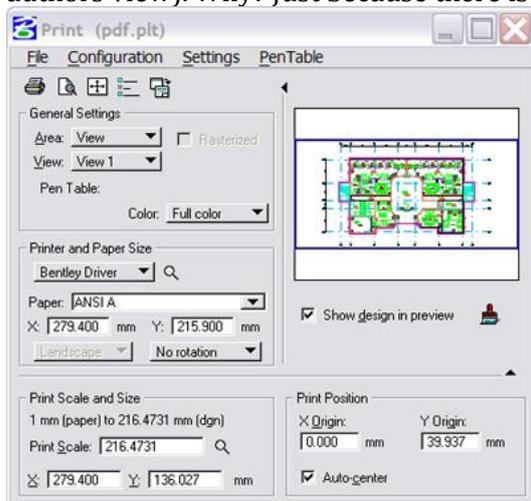


Figure 20: Misaligned, inconsistent and cluttered layout.

Figure 21: Aligned and consistent interface.

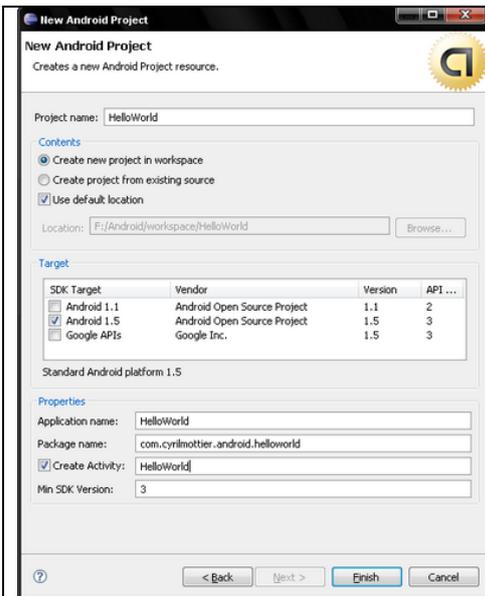


Figure 22: uncluttered interface.

❖ Advantages

- Simple formula, straightforward to implement.

❖ Drawbacks

- A dense interface is not necessarily unaesthetic! Fig.21b is not necessarily less aesthetic than Fig.20a.

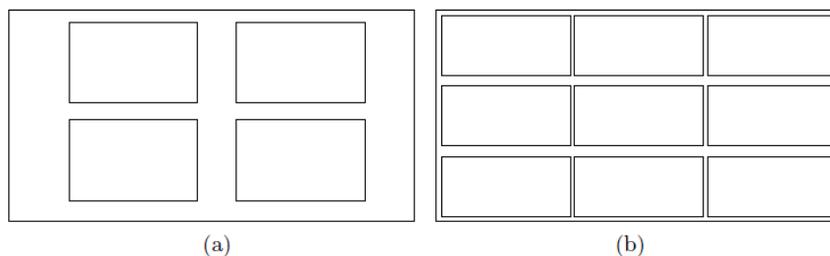


Figure 23: Two versions of screens in a density study: (a) a good version, (b) a bad version.

III.5.2 Balance

Aesthetics Metric

❖ Metric Name

Balance

❖ References

- [5] D.C.L. Ngo and J.G. Byrne, "Another Look at a Model for Evaluating Interface Aesthetics", International Journal of Applied Mathematics and Computer Science (European publisher), Vol. 11, No. 2, 515-535
- [16] Vanderdonckt, J. "Visual Design Methods in Interactive Applications", 2002

DEFINITION

Balance can be defined as the distribution of the optical weight in a picture. The optical weight refers to the perception that some objects appear heavier than others. Larger objects are heavier, whereas small objects are lighter. Balance in screen design is achieved by providing an equal weight of screen elements, left and right, top and bottom. [5]

Balance is a search for equilibrium along a vertical or horizontal axis in the layout. If a weight is attached to every IO, balance requires that the sum of IO weights on each hand of the axis remains similar (fig.22a). Balance is justified by the human perception and intense need for it in visual layouts. Balance is a technique in which there exists a gravity centre located on a vertical or horizontal axis. The opposite of balance is instability where IO are not distributed equally on each hand of the axis (fig. 22b). "They seem ready to topple over." [16]

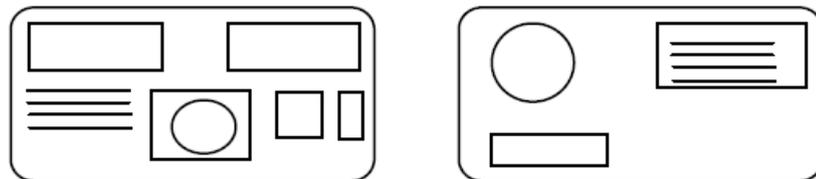


Figure 24: Balanced and unbalanced layouts

- *Comments* : Balance implies equilibrium. So the *equilibrium* metric [5] can be discarded.
- *Importance/Weight*: high (Some authors [16] say Balance is the most important metric to achieve.)

❖ Formula

This is the formula provided by Ngo[5]:

$$BM = 1 - \frac{|BM_{\text{vertical}}| + |BM_{\text{horizontal}}|}{2} \in [0, 1].$$

Here BM_{vertical} and $BM_{\text{horizontal}}$ are respectively the vertical and horizontal balances with

$$BM_{\text{vertical}} = \frac{w_L - w_R}{\max(|w_L|, |w_R|)},$$

$$BM_{\text{horizontal}} = \frac{w_T - w_B}{\max(|w_T|, |w_B|)},$$

where

$$w_j = \sum_i^{n_j} a_{ij} d_{ij}, \quad j = L, R, T, B,$$

L, R, T and B stand for left, right, top and bottom, respectively, a_{ij} is the area of object i on side j , d_{ij} is the distance between the central lines of the object and the frame, n_j is the total number of objects on the side.

❖ Motivation

The interface below is beautiful but unstable because it's not symmetric:



Figure 25: Unstable interface.

The same interface looks better when made symmetric. Symmetry implies balance.



Figure 26: Same interface; edited in Microsoft Paint.

❖ Advantages

- ...

❖ Drawbacks

- ...

III.5.3 The discarded metrics

Here are the metrics that we eliminated.

1. Equilibrium[2]
2. Sequence[2]
3. Cohesion[2]
4. Regularity[2]
5. Rhythm[2]
6. Horizontality[16]
7. Understatement/Exaggeration[16]
8. Neutrality/Accent[16]
9. Singularity/juxtaposition[16]
10. Negativity/Positivity[16]
11. Transparency/Opacity[16]
12. Consistency/Variation[16]
13. Predictability/Spontaneity[16]
14. Sequentiality/Randomness[16]
15. Continuity/Episodicity[16]
16. Sharpness/Diffusion[16]
17. Roundness/Angularity[16]
18. Stability/Stress[16]
19. Leveling/Sharpening[16]
20. Activeness/Passiveness[16]
21. Subtlety/Boldness[16]
22. Representation/Abstraction[16]
23. Realism/Distortion[16]
24. Flatness/Depth[16]
25. The family of Consistency [4] metrics and tools.

These metrics were discarded for the following reasons:

- Redundancy with respect to another metric that was already selected.
 - *Equilibrium[2]* w.r.t. *Balance*
 - *Cohesion[2]* w.r.t. *Unity*
 - *Regularity[2]* w.r.t. *Alignment*
 - *horizontality* w.r.t. *Proportion*
 - *Widget totals[4]* w.r.t. *Simplicity & Economy*
 - *Non-widget area[4]* w.r.t. *Density*
 - *Widget density[4]* w.r.t. *Density*
 - *Gridedness[4]* w.r.t. *Alignment*
 - *Area balances[4]* w.r.t. *Balance*
- Related to pure usability/ergonomics rather than aesthetics
 - *Concordance tools[4]*
 - *Consistency/Variation[16]*
 - *Predictability/Spontaneity[16]*

(However to some degree, consistency can be perceived as a pure aesthetic criterion but there were no formula proposed in the literature and devising one ourselves proved to be a non-trivial and time-consuming task)

- Require task analysis to be computed accurately.
 - *Understatement/Exaggeration[16]*

- Too complex to implement because they are rather based on a visual feeling and impression.
 - *Sharpness/Diffusion*[16]
 - *Roundness/Angularity*[16]
 - *Stability/Stress*[16]
 - *Leveling/Sharpening*[16]
 - *Activeness/Passiveness*[16]
 - *Subtlety/Boldness*[16]
 - *Representation/Abstraction*[16]
 - *Realism/Distortion*[16]
 - *Flatness/Depth*[16]
 - *Continuity/Episodicity*[16]
 - *Rhythm*[2]
- Irrelevance as aesthetic criterion!
 - *Sequence*[2]
 - *Sequentiality/Randomness*[16]
 - *Rhythm*[2]
 - *Transparency*[16]
 - *Negativity/Positivity*[16] (both are ok)
 - *Singularity/juxtaposition*[16]
 - *Neutrality/Accent*[16] (accent can and should be used to an important IO)
 - *Aspect Ratio*[4]
 - *Margins*[4]
 - *Distinct Typefaces*[4]
 - *Distinct Background Colors*[4]
 - *Distinct Foreground Colors*[4]

Chapter IV. Implementation

IV.1 Technologies used

IV.1.1 UsiXML

UsiXML[1] (USeR Interface eXtensible Markup Language) is an XML-based markup language for defining user interfaces, a User Interface Description Language aimed at describing user interfaces with various levels of details and abstractions, depending on the context of use. In other words you can specify a UI in terms of functionality (task analysis), object it manipulates, or, in a more concrete way, you can draw user interfaces. USIXML supports a family of user interfaces such as, but not limited to: device-independent, platform-independent, modality independent, and ultimately context-independent. USIXML allows specifying multiple models involved in user interface design such as: task, domain, presentation, dialog, and context of use, which is in turn decomposed into user, platform, and environment. These models are structured according to the four layers of the Cameleon framework: task & concepts, abstract user interface, concrete user interface, and final user interface (see fig. 27).

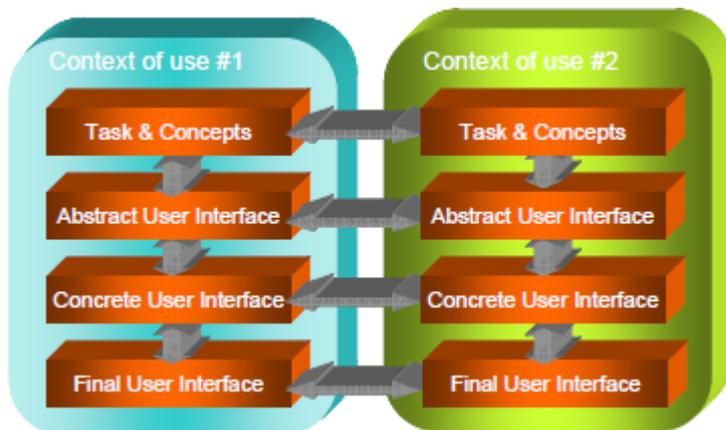


Figure 27: Structure of UsiXML; The four basic levels of the Cameleon reference framework.

In this thesis we consider only interfaces specified at the CUI (Concrete User Interface) level. And since the metrics that we need to compute require information that includes the coordinates of widgets, we decided, after deep consideration and advice from the supervisor and the teaching staff, to limit ourselves to the interfaces that are based on the *GridBagBox* layout, which includes coordinates of widgets.

IV.1.2 REST

REST stands for **R**epresentational **S**tate **T**ransfer. It relies on a stateless, client-server, cacheable communications protocol -- and in virtually all cases, the HTTP protocol is used. REST is *an architecture style* for designing networked applications. The idea is that, rather than using complex mechanisms such as CORBA, RPC or SOAP to connect between machines, simple HTTP is used to make calls between machines. (In many ways, the World Wide Web itself, based on HTTP, can be viewed as a REST-based architecture).

RESTful applications use HTTP requests to post data (create and/or update), read data (e.g., make queries), and delete data. Thus, REST uses HTTP for all four CRUD (Create/Read/Update/Delete) operations.

REST is a lightweight alternative to mechanisms like RPC (Remote Procedure Calls) and Web

Services (SOAP, WSDL, et al.).

REST is not a "standard". There will never be a W3C recommendation for REST, for example. And while there are REST programming frameworks, working with REST is so simple that you can often "roll your own" with standard library features in languages like Perl, Java, or C#.

IV.1.3 DOM

The Document Object Model (DOM) is a cross-platform and language-independent convention for representing and interacting with objects in HTML, XHTML and XML documents.

The DOM presents an XML document as a tree-structure (see example fig.28).

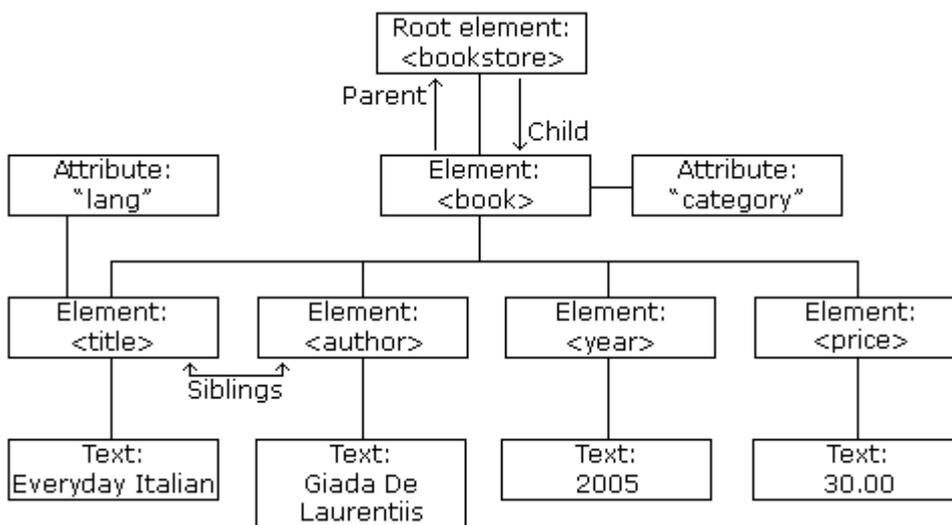


Figure 28: XML DOM Example.

Aspects of the DOM (such as its "Elements") may be addressed and manipulated within the syntax of the programming language in use. The public interface of a DOM is specified in its application programming interface (API).

IV.1.4 Java

➤ JAX-RS

JAX-RS ("Java API for RESTful Web Services") is a Java API that provides support in creating web services according to the Representational State Transfer (REST) architectural style. JAX-RS uses annotations, introduced in Java SE 5, to simplify the development and deployment of web service clients and endpoints.

➤ Jersey

Jersey is Sun's production quality reference implementation for JSR 311: JAX-RS. Jersey implements support for the annotations defined in JSR-311, making it easy for developers to build RESTful web services with Java and the Java JVM.

➤ JAXP

The Java API for XML Processing (JAXP) allows applications to parse and transform XML documents using an API that is independent of any particular XML processor implementation. Through a plug-in scheme, developers may change the XML processor implementations without impacting their applications. JAXP 1.1 supports the following

standards:

- DOM Level 2
- SAX version 2.0
- XSLT 1.0

IV.2 Architecture

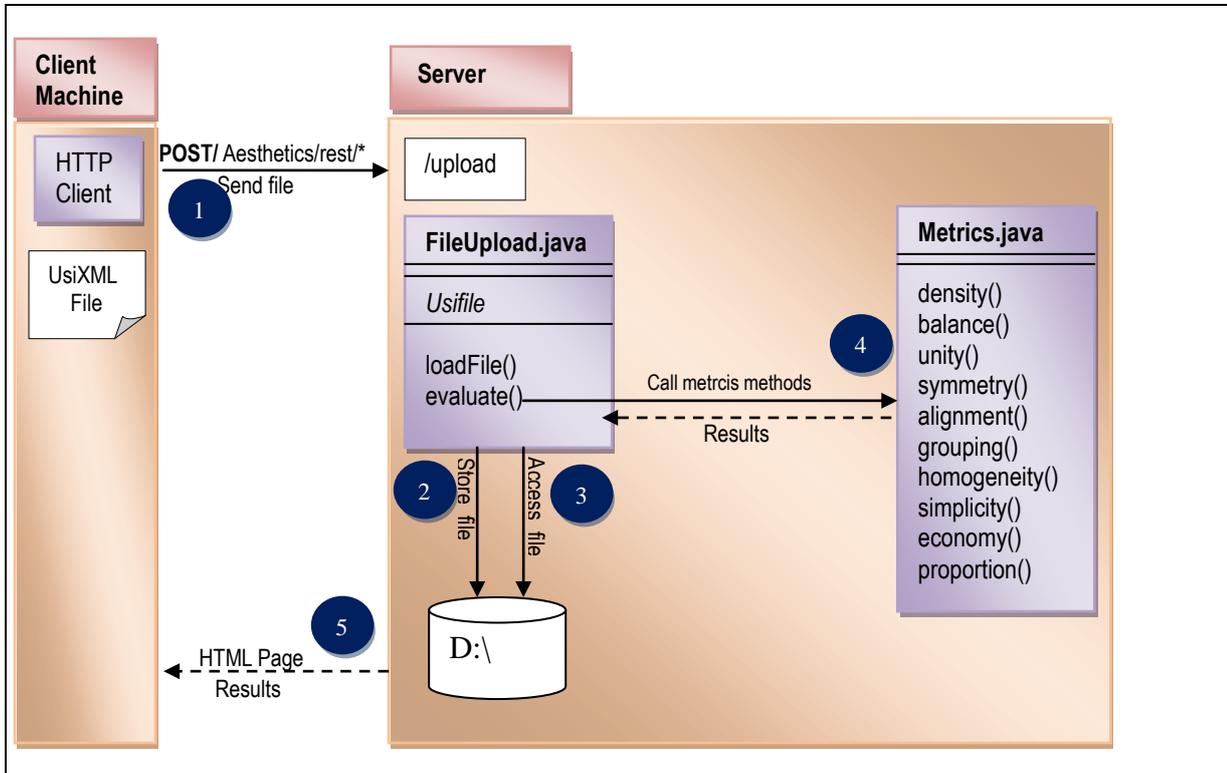


Diagram 5: Typical scenario of a client requesting an assessment of a UsiXML GUI.

Diagram 5 describes the architecture of the application.

The automated evaluator is implemented in the form of a REST web application that is available under the following path: “*http://any_domain:port/AesthetiX/rest/upload*”.

The client is required to provide a well-formed UsiXML file, describing a Graphical User Interface based on a *GridBagBox* layout [1].

The server receives the uploaded file and stores it in the local memory (method *loadFile()*) and then opens the file and calls the methods from the class *Metrics.java* to evaluate each metric on the interface.

The server returns the results for each metric as well as the total score for the interface in the form of a HTML web page.

If the file provided by the client doesn’t meet the requirements, the result is one of the following.

- If the file is a well-formed XML file, the webpage that is returned is empty and the user is invited to check that the file he provided is a well-formed UsiXML file.
- If the file is not a well-formed XML file, an error message is returned alongside the description of the error (i.e. the exception generated by java)

The typical scenario of diagram 5 is replicated in appendix E with screenshots of the actual java code.

Chapter V. Tests

In this section we are going to describe the experiment that we conducted in order to evaluate the effectiveness of using the proposed measures to assess interfaces.

V.1 Screen Selection and Automated Evaluation

V.1.1 The screens

We first needed to find an “interesting” set of interfaces developed in UsiXML. Having scoured the available resources and knocked on the doors of the teaching staff, we were unable to find such a set. So we decided to make our own set of screens. However this task also proved to be rather difficult: the tools available to edit and view UsiXML interfaces are not as advanced as we had expected. (We thought it would be as easy as creating web pages in HTML). So we had to rethink our strategy: we decided to select basic dialog boxes with interesting characteristics; and then use GraphiXML[1] to sort of make an equivalent “skeleton” thereof in UsiXML (i.e. a manual copy having the same layout). Here are the interfaces that we selected (figures 29, 31, 33, 35, 37) along with the corresponding UsiXML copy (figures 30, 32, 34, 36, 38) and the output from the program (table 1 through 5):

1. First screen: Android

This interface (fig.29) is the best in our view. It is the exemplary good screen in our selection. It should receive good ratings both from the program and the users and should outrank the other interfaces.

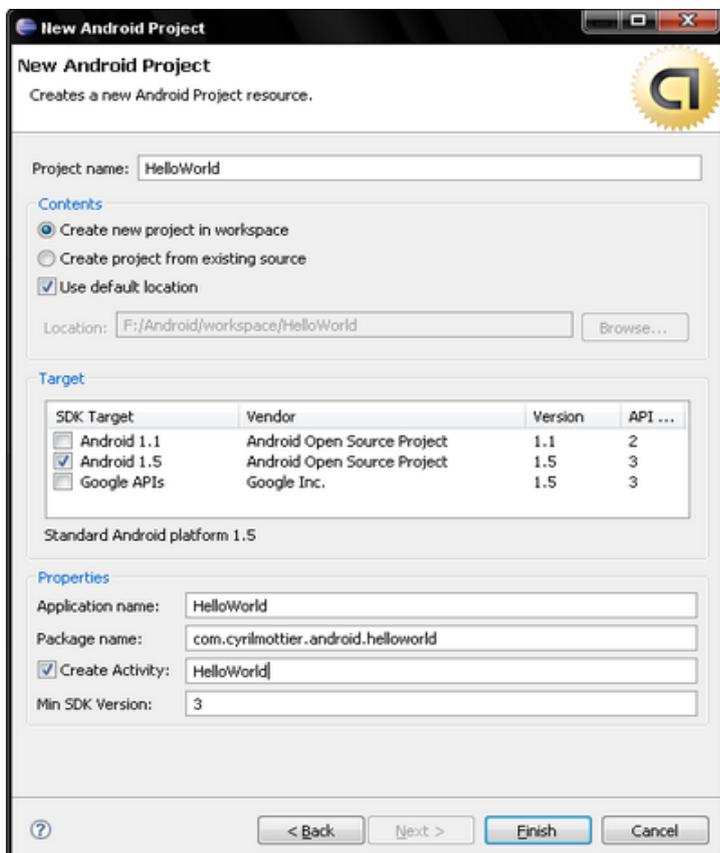


Figure 29: good screen that should receive good ratings from the program and the users.

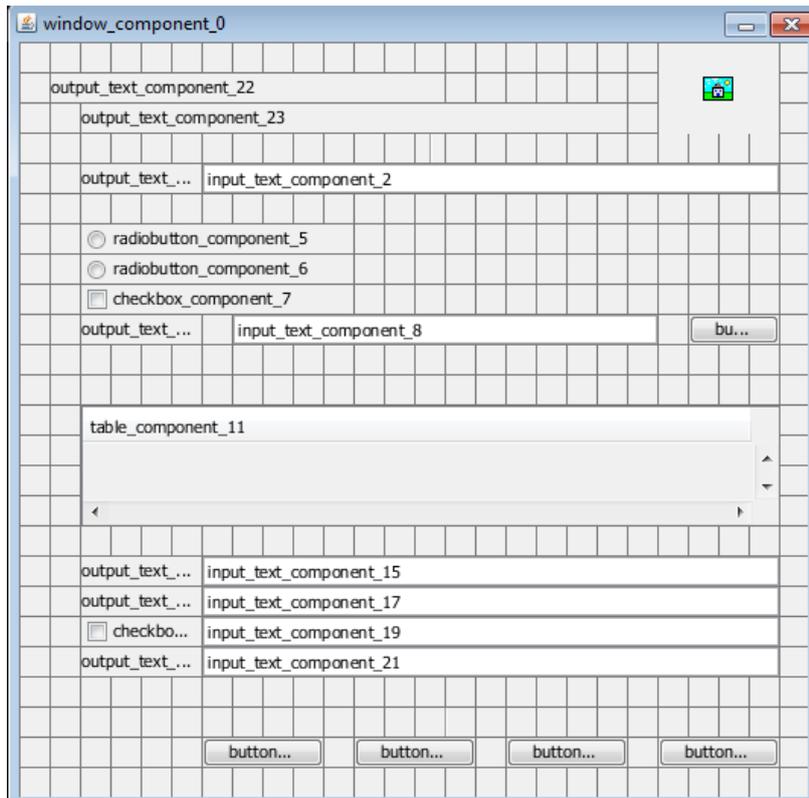


Figure 30: UsiXML manual copy of the interface “Android”.

Table 1: Output of the program for Android.

Metrics	Ratings ($\in [0,1]$)	Comments
Density	0.44153845	ok (this score makes sense)
Balance	0.95717	ok
Unity	0.42364648	Makes sens w.r.t. formula: the 'outputtext' and 'inputtext' are far from each other.
Symmetry	0.85290664	ok
Alignment	0.92692304	ok
Grouping	0.6252952	ok
Repartition	0.0102040816326	Low value because the formula is based on the number of objects, not their area.
Simplicity	0.04918033	Low value because the formula is too simplistic (see form 6 of Appendix C)
Economy	0.10526316	Low value because the formula is too simplistic (see form 7 of Appendix C)
Proportion	0.627174	ok
Overall Score	0.710447543558	

2. Second screen: Ubuntu Firewall

We selected this interface (fig.31) because it has an obvious balance/symmetry problem. The checkboxes and radio boxes on the left have no counterparts on the other side of the screen. And the icon is not centered. However it still looks beautiful because of its low density.



Figure 31: Unbalanced screen.

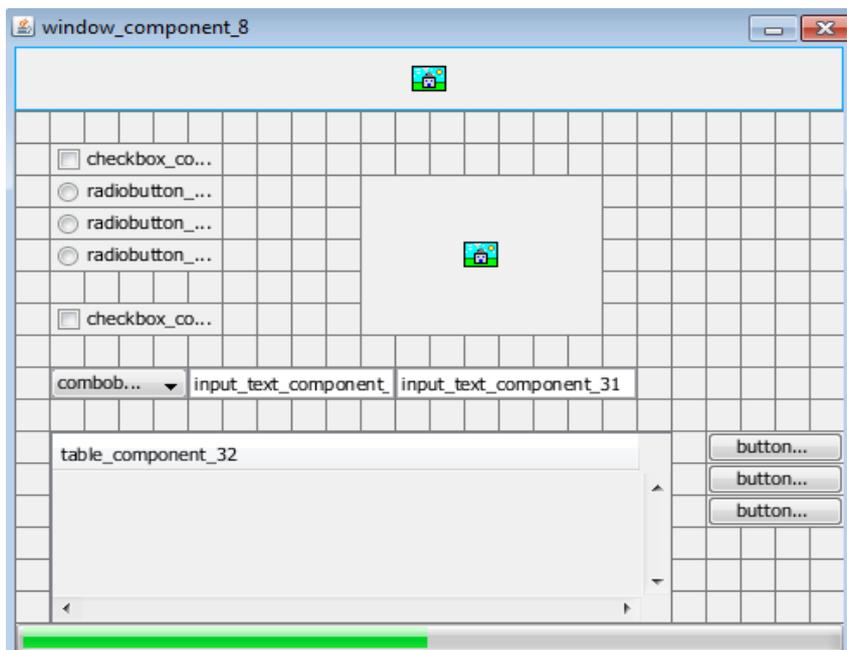


Figure 32: UsiXML copy of “Ubuntu firewall”.

Table 2: Output of the program for Ubuntu firewall.

Metrics	Ratings ($\in [0,1]$)	Comments
Density	0.4100877	ok (this score makes sense)
Balance	0.6110871	ok
Unity	0.26666665	ok
Symmetry	0.7713198	ok
Alignment	0.84880954	ok
Grouping	0.6477578	ok

Repartition	0.0075	Low value because the formula is based on the number of objects, not their area.
Simplicity	0.069767445	Low value because the formula is too simplistic (see form 6 of Appendix C)
Economy	0.125	Low value because the formula is too simplistic (see form 7 of Appendix C)
Proportion	0.63293254	ok
Overall Score	<i>0.5510024463295</i>	

3. Third screen: Ubuntu Firewall Balanced

We rectified the balance problem in Microsoft Paint by centering the icon and dragging the buttons on the left upwards to make the weight even on both sides. Thus, we expected this interface (fig.33) to receive better evaluations than the previous one, both from the program and the users.



Figure 33: Same interface as in fig.31 with better balance.

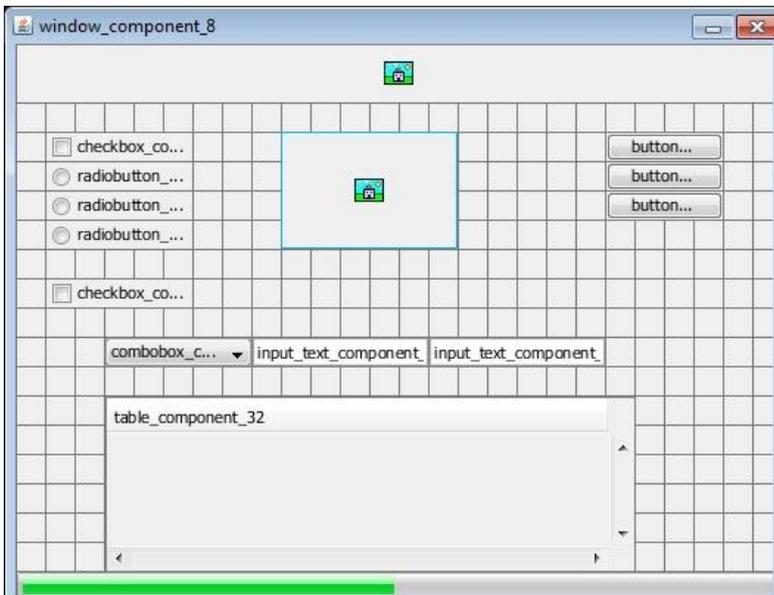


Figure 34: UsiXML copy of Ubuntu firewall.

Table 3: Output of the program for the balanced version of “Ubuntu firewall”.

Metrics	Ratings $\in [0,1]$	Comments
Density	0.4505263	ok
Balance	0.63308024	ok (but we expected the difference with the previous screen to be bigger)
Unity	0.28333333	ok
Symmetry	0.67637646	We expected this screen to have better symmetry than the previous one.

Alignment	0.87142855	ok
Grouping	0.7198327	ok
Repartition	0.00178571428	Low value because the formula is based on the number of objects, not their area.
Simplicity	0.07317073	Low value because the formula is too simplistic (see form 6 of Appendix C)
Economy	0.13333334	Low value because the formula is too simplistic (see form 7 of Appendix C)
Proportion	0.6602882	ok
Overall Score	0.56349461321106	

4. Fourth screen: Printer Dialog box

We selected this interface (fig.35) because it has an obvious misalignment problem. This is due to the inconsistent sizes and positions of the labels, input fields and combo boxes. It is also very cluttered and looks quite messy. We expect this interface to receive the lowest rating both from the program and the users.

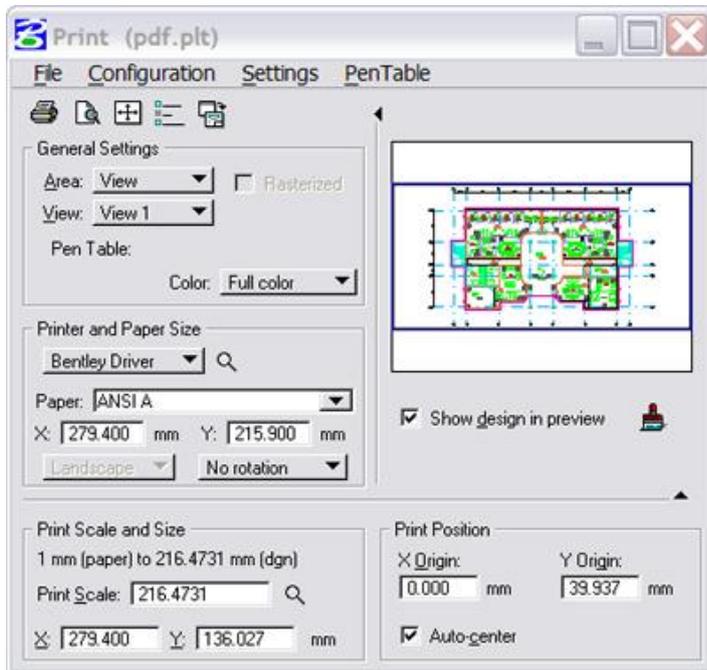


Figure 35: Misaligned interface

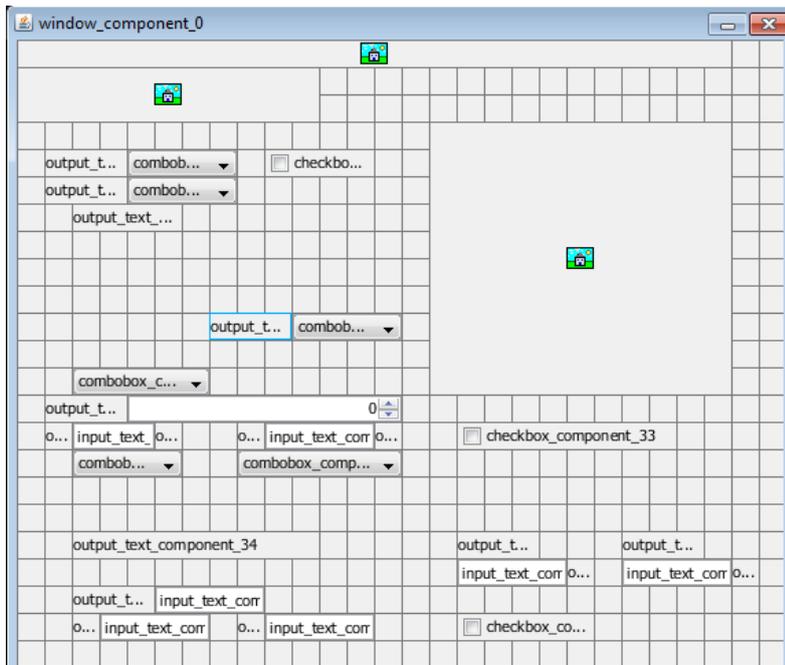


Figure 36: UsiXML copy of “Printer dialog box”.

Table 4: Output of the program for “Printer dialog box”.

Metrics	Ratings ($\in [0,1]$)	Comments
Density	0.56055903	ok
Balance	0.6815569	ok
Unity	0.43742982	ok
Symmetry	0.793388	ok
Alignment	0.93862605	Rating should be bad. Formula needs to be improved (see form 3; Appendix C)
Grouping	0.4875769	ok
Repartition	0.0000925545	Low value because the formula is based on the number of objects, not their area.
Simplicity	0.037037037	Low value because the formula is too simplistic (see form 6 of Appendix C)
Economy	0.11764706	Low value because the formula is too simplistic (see form 7 of Appendix C)
Proportion	0.7240957	ok
Overall Score	0.6274979310755	

5. Fifth screen: Printer Dialog Box Aligned

We rectified the alignment problem in Microsoft Paint by shortening some labels, stretching some input fields and combo boxes and moving some of them to another position. Thus, we expected this interface to receive better evaluations than the previous one, both from the program and the users

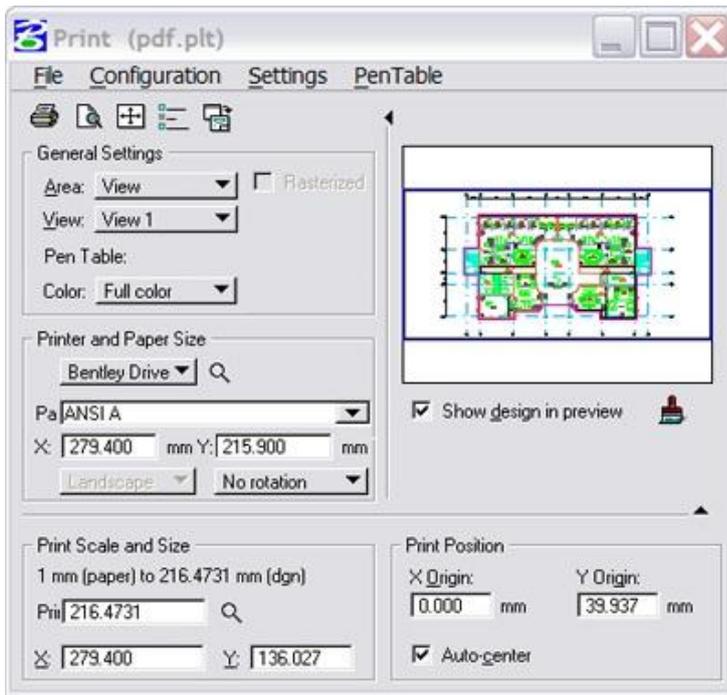


Figure 37: Same interface as in fig. 35 with better alignment.

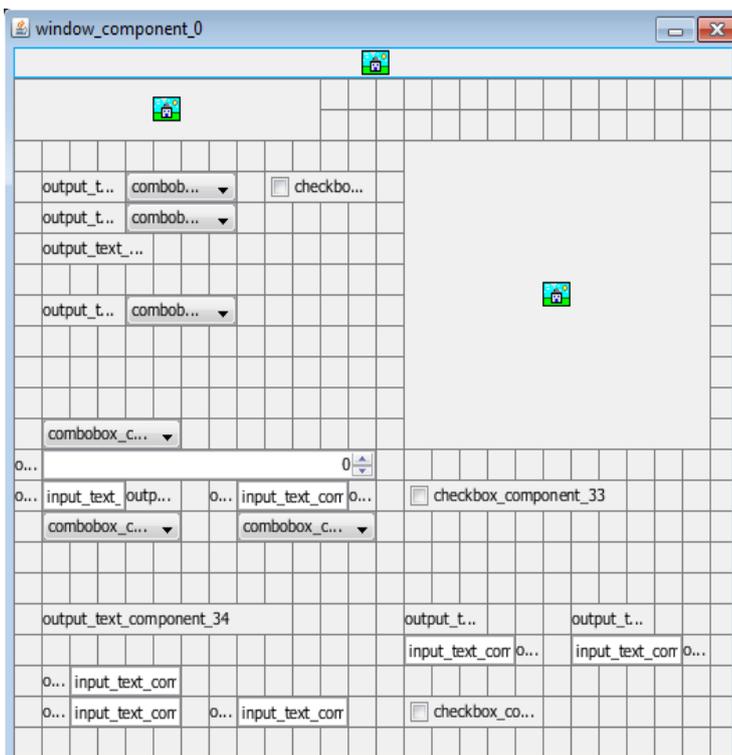


Figure 38: UsiXML copy of the interface in fig. 37.

Table 5: Output of the program for the aligned version of “Printer dialog box”.

Metrics	Ratings ($\in [0,1]$)	Comments
Density	0.5317726	ok
Balance	0.6882392	ok
Unity	0.43952915	ok
Symmetry	0.78966725	ok
Alignment	0.9457583	ok
Grouping	0.5236845	ok

Repartition	0.0000925545	Low value because the formula is based on the number of objects, not their area.
Simplicity	0.03797468	Low value because the formula is too simplistic (see form 6 of Appendix C)
Economy	0.11764706	Low value because the formula is too simplistic (see form 7 of Appendix C)
Proportion	0.6718385	ok
Overall Score	0.6259701198984575	

V.1.2 Discussion

The results (tables 1 to 5) are consistent with respect to the formulas; however some formulas' accuracy can be questioned:

- The formula that we defined (and refined several times) for *alignment* wasn't able to make a difference between "printer dialog box" (fig.35), or rather its UsiXML copy (fig.36), and the aligned version of printer dialog box. Will the users make a difference?
- The formulas for *simplicity* and *economy* are too simple (see forms 6 and 7 of Appendix C). But comparatively (when we compare the results for each screen), their scores are rather consistent.
- *Proportion* is comparatively consistent and accurate.
- *Repartition* gives very low values because it is based on the number of objects on each quadrant instead of the areas of the objects. However the results are comparatively consistent.
- The formula that we defined for *grouping* gives rather accurate results.
- *Unity* is comparatively consistent and accurate.
- *Balance* is comparatively consistent and accurate.
- *Density* is rather consistent and accurate (with regard to the UsiXML manually copies the interfaces).
- *Symmetry* seems comparatively consistent and accurate too; however we expected the balanced version of Ubuntu firewall to receive a better value than the non-balanced one.

Notes:

- The evaluation of the interfaces is done on their UsiXML copy (fig. 30, 32, 34, 36, 38). This approach introduces a slight bias of its own with respect to the values we would get on the original screen.
- The overall score is a weighted average. The weights are assigned to each metric based on its degree of importance. Balance, the most important one, received a weight of .35; density and alignment received .15; unity and symmetry received .10; and the rest received .03.

V.2 Subjective Ratings by Users

V.2.1 The experiment

We conducted an experiment to evaluate the effectiveness of using the proposed measures and tool to assess interfaces. As we said in Chapter III, sections 1 and 2.3, we needed a universal and absolutely objective observer to validate our tool. And the closest instance of such an observer is a heterogeneous population of human evaluators (the "God-crowd"). So we invited 16 people of various ethnic backgrounds, ages and sex (see table 6) and convinced them to participate in our experiment. Only six of them were familiar with the fields of Information Technology. The rest did not even know what we meant by "interface". Three of them were familiar with visual

design/interface design concepts.

Total	16
Sex	
Male	8
Female	8
Ethnicity	
Western Europe	4
North America	1
Latin America	2
Northern Africa	3
Sub-Saharan Africa	5
Far East	1
Age	
< 20	2
20-24	6
25-29	4
30-40	1
40-50	2
>50	1

Table 6: Statistics on participants.

The participants were presented the five interfaces from the previous section and were asked to rate them on a low-medium-high scale regarding how aesthetic (or beautiful) it was. They were given only five minutes to complete the form in appendix D because, on the one hand, we wanted their judgment to be based on first impressions and on the other hand we wanted to save their time. We also purposely made this experiment simple for two reasons: first, we wanted the experiment to be easy to understand for people from various backgrounds and education fields and levels. And then, we wanted to arouse their interest by claiming that we had a tool that (we believed) could measure beauty and asking them to simply

“evaluate the beauty of interfaces” so we could compare their opinions to our program’s results. Unlike what we did in our previous work (see appendix B), we did not want to explain the metrics and design concepts to the participants; in this way we were able to attract them as we planned in our methodology (see section 1, chapter III).

The average for the ratings was calculated to show the relative aesthetics of the five screens, according to the viewers (see table 7).

Interface	Android	Ubuntu	Balanced Ubuntu	Printer	Aligned Printer
Score	2,625	2,1875	2,4375	1,4375	1,5

Table 7: Averages of subjective ratings. (low=1, medium=2, high=3).

V.2.2 Discussion

Rank	Program	Users
1	Android	Android
2	Printer Aligned Printer	Balanced Ubuntu
3		Ubuntu
4	Balanced Ubuntu	Aligned Printer
5	Ubuntu	Printer

Table 8: Rankings of the interfaces by the program and by the users.

We now have the evaluation of an objective observer, the “god-crowd”. Let us now confront the results from the interviews with those from our program. (Table 8 compares the rankings):

- As expected, Android comes first in both rankings.
- The Ubuntu interfaces are rated better than the Printers; this was intended and expected when selecting our interfaces. The printers are very cluttered and messy while the Ubuntu look nice because of the good use of space. We hoped that *density*, *simplicity* and *economy*

would cause the program to give better evaluations to the Ubuntu but they didn't. We think that this is due to the fact that we are not evaluating the actual interfaces but their manual copy; because the metrics seem rather accurate (except for *alignment*).

- We thought that the problem with the printer dialog box was misalignment; so we tried to fix the problem and expected the aligned version to rank better than the original version. However, the users couldn't tell the difference between those two interfaces and we were often asked during the interviews "why the same interface appeared twice" in our screenshots. Well that's good news! Because the program couldn't tell the difference either.
- As expected, the balanced version of Ubuntu received a better evaluation than the original one in both rankings. Unlike what was expected, the difference is very small! But it's small in both rankings; that's good news. When shown the Ubuntu interfaces the users couldn't tell the difference until we explained it to them.

Conclusions

In this thesis, we first demonstrated the importance of aesthetics in HCI. Not only does aesthetics increase perceived usability but it also contributes to customer/user satisfaction. However, this aspect of the interaction has been widely neglected until relatively recently by usability testers;

If aesthetics considerations should be included in these tests, the problem then becomes how to evaluate aesthetics. Is it even possible to evaluate aesthetics?

To answer these questions, we scoured the philosophical literature on the subject and came to the following conclusions:

Roughly speaking there are three schools of thought: *subjectivism*, *relativism* and *objectivism*. The philosophy that we adopt determines how we can attempt to measure beauty.

If we (as usability testers or designers) believe in *subjectivism* then we ought to conduct brain studies and research, for example, in order to find out what in a person's brain causes them to have the particular preferences that they have. Then we could for example survey the target population through questionnaires to determine their psychological traits, so we can design our tools accordingly. If on the other hand, we are *relativists* then a sociological study is a must!

In this thesis we adopted the *objectivist* point of view. Remember that in an *objectivist* worldview, aesthetic qualities belong to the object independently of the subjects. This means that we can implement a tool that recognizes those aesthetic features! Every subject/observer will acknowledge them as well, provided that they can perceive them. Their vision can be hindered by various psychological "veils". So the closest instance of an "unveiled observer" is a large number of individuals. That's why we tested our tool against a set of users from various backgrounds.

Statistically speaking there is a shortcoming that we can already point out: we do not know how many individuals or what type of individuals we should include in our sample; maybe one Mister Joe Public (a randomly selected individual) is enough to test our tool! He might have no psychological "veils" and thus could be our reference. On the other hand, one million individuals could all have the same "veils" and cause us to incorrectly (in)validate the results from the program. *Note*: in an *objectivist* worldview, the culture (ethnicity, context, gender, age ...) of the participants might shape the "veil" that they have.

From the discussions in the previous chapter (sections V.1.2 and V.2.2) we it appears that the users confirmed what we already expected when selecting the interfaces to test (even though, admittedly they couldn't see any significant difference between the improved versions of screens and their original ones). The program, on the other side, seems rather accurate. Except for *alignment* (whose accuracy we doubt), the metric's set seems accurate and was able to rank Android as first, to point out close resemblance of interfaces we expected to be significantly different. We believe that the difference in ranking is not due to the program but to the fact that the program is not evaluating the same thing as the users. The program is evaluating the manual UsiXML copy of the interfaces. And like we said, the results it gives seem consistent with respect to those copies.

Consequently, the first thing that future work should include is a test of our program on real UsiXML interfaces.

What have we achieved then? What did we prove in this thesis, if anything?

If we assume that the "God-crowd" was really an objective observer, then we proved that some basic features of aesthetics (e.g. *balance*) are essential and can be measured. Those features can be included in a tool to help designers and usability testers do their job. However, there is more to aesthetic experience than the basic features that we considered in our basic interfaces (see chapter 3 section III.3.4). And we doubt that in its present state the developed tool can be of any help to

professional design and usability testing. There is a need for much more advanced metrics. And those metrics should include task analysis in accordance with the principle of *aesthetic duality* (see Chapter 3 section III.2).

If, on the other hand, we assume that the God-crowd's judgment wasn't objective (because of some cultural or psychological veils shared by all the participants), then the author has the same veils and the metrics set that he selected only reflects the veils that we share. Hence the concordance in evaluations.

References

- [1] UsiXML website: <http://www.usixml.org>
- [2] D.C.L. Ngo, L.S. Teo, J. G. Byrne *A Mathematical Theory of Interface Aesthetics* available at <http://www.mi.sanu.ac.rs/vismath/ngo/index.html>
- [3] Mahajan, R. and Shneiderman, B. (1995), A family of user interface consistency checking tools: Design analysis of SHERLOCK, *Proc. of NASA Twentieth Annual Software Engineering Workshop, SEL-95-004*, NASA Pub., 169-188
- [4] Mahajan, R. and Shneiderman, B. *Visual & Textual Consistency Checking Tools for Graphical User Interfaces*, 1996.
- [5] D.C.L. Ngo and J.G. Byrne, "Another Look at a Model for Evaluating Interface Aesthetics", *International Journal of Applied Mathematics and Computer Science* (European publisher), Vol. 11, No. 2, 515-535
- [6] D.C.L. Ngo, A. Samsudin and R. Abdullah, "Aesthetic Measures for Assessing Graphic Screens", *Journal of Information Science and Engineering* (Taiwanese publisher), Vol. 16, No. 1, 97–116, 2000.
- [7] Tractinsky, N. (1997) Aesthetics and Apparent Usability: Empirically Assessing Cultural and Methodological Issues, *CHI 97 Conference Proceedings*, Atlanta, March 22-27, 1997), ACM, New York, pp. 115-122
- [8] Tractinsky, N., Shoval-Katz A. and Ikar, D. (2000) what is Beautiful is Usable *Interacting with Computers*, 13(2): 127-145
- [9] Lavie, T. and Tractinsky, N. (2004) Assessing Dimensions of Perceived Visual Aesthetics of Web Sites, *International Journal of Human-Computer Studies*, 60(3):269-298
- [10] Tractinsky, N. (2004) Towards the Study of Aesthetics in Information Technology, *25th Annual International Conference on Information Systems*, Washington, DC, December 12-15, pp. 771-780
- [11] Tractinsky, N., Cokhavi, A., and Kirschenbaum, M. (2004) Using Ratings and Response Latencies to Evaluate the Consistency of Immediate Aesthetic Perceptions of Web Pages *Third Annual Workshop on HCI Research in MIS*, Washington, D.C., December 10-11, 2004, pp. 40-44
- [12] Tractinsky, N. and Zmiri, D. (2005) Exploring Attributes of Skins as Potential Antecedents of Emotion in HCI in Fishwick, P. (ed.) *Aesthetic Computing*, MIT Press (in press).
- [13] Avraham Parush, Ronen Nadir, Avraham Shtub, *Evaluating the Layout of Graphical User Interface Screens: Validation of a Numerical Computerized Model*, *International Journal of Human-Computer Interaction*, Volume 10, Issue 4 December 1998 , pages 343 – 360
- [14] Gary Perlman, *An Axiomatic Model of Information Presentation*, Paper originally published in *Proceedings of the Human Factors Society 31st Annual Meeting*, 1987, 1229-1233, Santa Monica, CA: Human Factors Society

- [15] B. Shneiderman, R. Chimera and N. Jog, R. Stimart, D. White, *Evaluating Spatial and Textual Style of Displays*, May 1995
- [16] Vanderdonckt, J. “*Visual Design Methods in Interactive Applications*”, 2002
- [17] Vanderdonckt, J. and Gillo, X. (1994), “Visual techniques for traditional and multimedia layouts”, In Catarci, T., Costabile, M., Levialdi, S. and Santucci, G. (Editors), *Proc. Advanced Visual Interfaces Conference '94*, ACM Press, New York, 95-104.
- [18] Carrie A. Lee, *HCI & Aesthetics: The Future of User Interface Design* available at <http://www.carrieelee.net/pdfs/HCIAesthetics.pdf>
- [19] K. Hartmann , T. Götzelmann , K. Ali and T. Strothotte, *Metrics for Functional and Aesthetic Label Layouts*, Springer Berlin / Heidelberg, 2005, pages 115-126
- [20] J. Hartmann , A. Sutcliffe , A. De Angeli, *Towards a Theory of User Judgment of Aesthetics and User Interface Quality*, ACM Transactions on Computer-Human Interaction, Vol.15, No.4, Article 15, November 2008.
- [21] D. C. L. Ngo, L. S. Teo and John G. Byrne, *Modelling interface aesthetics*, Information Sciences, Volume 152, June 2003, Pages 25-46
- [22] L.L. Constantine, "Visual Coherence and Usability: A Cohesion Metric for Assessing the Quality of Dialogue and Screen Designs," *ozchi*, pp.0115, 6th Australian Conference on Computer-Human Interaction (OZCHI '96), 1996
- [23] Michael Bauerly, Yili Liu, *Effects of Symmetry and Number of Compositional Elements on Interface and Design Aesthetics*, International Journal of Human-Computer Interaction, 1532-7590, Volume 24, Issue 3, 2008, Pages 275 – 287
- [24] Hassenzahl, M. 2004, *The interplay of beauty, goodness and usability in interactive products. Human-Comput. Interact.* 19, 4, 319-349.
- [25] S. S. Reilly and J. W. Roach, “Improved visual design for graphics display,” *IEEE Computer Graphics and Applications*, Vol. 4, No. 2, 1984, pp. 42-51.
- [26] E. Tjalve, *A Short Course in Industrial Design*, Newnes-Butterworths, Boston, 1979, pp. 141-173 [20.1, pp. 143-144].
- [27] O. Naukkarinen, *Why beauty still cannot be measured*, September 2010. Available at <http://www.contempaesthetics.org/newvolume/pages/article.php?articleID=603>
- [28] Rantalla, Veikko." *Environmental Experience: Beyond Aesthetic Subjectivism and Objectivism.*" Available at <http://www.lancs.ac.uk/users/philosophy/resources%20rtf%20files/rantalla.pdf>
- [29] *Web service*, From Wikipedia, the free encyclopedia. Wikimedia Foundation Inc. (http://en.wikipedia.org/wiki/Web_service)
- [30] *Golden ratio*, From Wikipedia: The Free Encyclopedia. Wikimedia Foundation Inc.
- [31] Nirmaldasan. 2004. Measuring aesthetics. *The Hindu* (Feb. 23). Available at

<http://www.hindu.com/edu/2004/02/23/stories/2004022300160400.htm>.

[32] A biography of George David Birkhoff can be found at <http://www-gap.dcs.st-and.ac.uk/~history/Mathematicians/Birkhoff.html>.

[33] Helen Balinsky, *Evaluating interface aesthetics: a measure of symmetry*, February 2006, Media Technology Laboratory, HP Laboratories Bristol.

[34] A. Vande Moere, N. Cawthon & A. Lau, *The Evaluation of Aesthetic Quality in Data Visualization*, 2006.

[35] S. J. Harrington, J. Fernando Naveda, R. P. Jones, *Aesthetic Measures for Automated Document Layout*, Proceedings of the 2004 ACM symposium on Document engineering, New York, NY, USA ©2004.

[36] Braun, C., Gruendl, M., Marberger, C. & Scherber, C. (2001). Beautycheck - Ursachen und Folgen von Attraktivitaet. Report. [pdf-document]. Available from: <http://www.beautycheck.de/english/bericht/bericht.htm>

[37] Pamela M. Pallett, Stephen Link, Kang Lee, *New "golden" ratios for facial beauty*, October 2009.

[38] Florian Hoenig, *Defining Computational Aesthetics*, 2005.

[39] J. M. Zain, M. Tey, and Y. Goh, *Probing a Self-Developed Aesthetics Measurement Application (SDA) in Measuring Aesthetics of Mandarin Learning Web Page Interfaces*, IJCSNS International Journal of Computer Science and Network Security, VOL. 8 No. 1, January 2008.

[40] Maryam Mirdehghani, and S. Amirhassan Monadjemi, *Web Pages Aesthetic Evaluation Using Low-Level Visual Features*, 2009.

[41] S.O. Hansson, *Aesthetic Functionalism*, October 2005. Available at <http://www.contempaesthetics.org/newvolume/pages/article.php?articleID=324>

[42] Russell, Bertrand (1919). "The Study of Mathematics". *Mysticism and Logic: And Other Essays*. Longman. p. 60.

[43] Bastien, J.M.C., Scapin, D. (1993) *Ergonomic Criteria for the Evaluation of Human-Computer interfaces*. Institut National de recherche en informatique et en automatique, France.

Appendix A: Metric Description Template.

Aesthetics Metric

❖ Metric Name

❖ References

DEFINITION

❖ Formula

❖ Motivation

❖ Advantages

❖ Drawbacks

Appendix B: A study of Vanderdonckt and Gillo's « Visual Techniques ».

Mesures esthétiques d'interfaces graphiques

{adrien.bourgeois, diego.dursel, simon.goffin, patrick.mbenza, jawira.portugal}
@student.uclouvain.be

23 juin 2011

1 Résumé

Dans ce travail, nous confrontons des appréciations subjectives d'interfaces avec leur évaluation effectuée à partir des "Visual Techniques" définis par Vanderdonckt et Gillo [2]. Pour cela nous avons interrogé un échantillon de 14 personnes. Ce travail met en évidence les contradictions qu'il peut parfois y avoir entre des conclusions tirées à partir des évaluations des critères et des conclusions tirées à partir d'évaluations subjectives.

2 Mots-clés

Critères d'esthétique, corrélation de Pearson, facteurs d'influence

3 Mots-clés (ACM classification)

H.5.1 [Information interfaces and presentation] : Multimedia Information Systems – Animations.

H5.2 [Information interfaces and presentation] : User Interfaces – User-centered design.

4 Termes généraux

Design, Experimentation, Human Factors, Measurement, Standardization, Theory, Verification.

5 Introduction

Il est universellement admis en marketing, en design et même en psychologie sociale que la beauté est importante. Elle est non seulement une qualité importante pour un produit (et contribue largement à son succès sur le marché), mais elle semble également influencer d'autres jugements concernant un objet [1]. À cet égard, plusieurs études ont suggéré une corrélation

entre la qualité esthétique d'une interface et son utilisabilité perçue (c-à-d Les belles choses donnent l'impression de mieux fonctionner [6]). Ce travail est basé sur les trente critères d'esthétique proposés par Vanderdonckt et Gillo [2] en vue d'aider les designers à créer des interfaces utilisateurs esthétiques. Nous les appellerons "critères d'esthétique" dans la suite de cet article; lequel a plusieurs objectifs :

Premièrement, nous nous proposons d'étudier le problème de la "subjectivité ou universalité" de l'esthétique. En effet nous savons de par la littérature que certains ont proposé des méthodes, principes et métriques en vue d'évaluer l'esthétique d'une interface graphique ([3], [4] et [5]) tandis que d'autres ont douté que l'esthétique puisse même être mesurée. Mais ici, nous nous focaliserons uniquement sur les trente critères d'esthétique : capturent-ils l'essence de l'esthétique (c-à-d permettent-ils de savoir si les utilisateurs vont aimer ou non une interface) ? Lesquels capturent le mieux l'essence de l'esthétique ? Et lesquels au contraire sont redondants et font double emploi ?

Ensuite, nous avons trouvé intéressant de chercher à savoir si ces critères d'esthétique sont "universels". Peuvent-ils être considéré comme reflétant simplement et universellement le bon sens ? Sont-ils absolus ou dépendent-ils de la culture, du sexe ou encore de l'âge ?

Et finalement, sachant que certaines études[1] (contredites par d'autres par ailleurs) ont suggéré qu'il y aurait une relation entre l'esthétique et l'utilisabilité perçue par les utilisateurs, nous voudrions savoir s'il y a une corrélation entre les trente critères d'esthétique et l'utilisabilité perçue.

Nous savons d'ors et déjà par [2] que ces critères sont difficilement applicables aujourd'hui parce que dans les interfaces modernes (spécialement les interfaces multimedia) les "layout grids" (voir [2]) ne sont plus applicables vu que le "layout" ne consiste plus en lignes verticales et horizontales mais en formes plus complexes (lignes obliques, formes convexes, surfaces planes, volumes,...). Dès lors, évaluer les critères sur de telles in-

terfaces n'est pas chose aisée.

Nous avons choisi ce sujet de travail parce que nous étions intéressé par l'esthétique des interfaces graphique et étions curieux d'en savoir plus sur les fameuses métriques d'esthétique ainsi que sur l'état actuel de la recherche dans ce domaine.

La suite de cet article est structurée de la manière suivante : la section suivante rapporte de manière non exhaustive l'état actuel de la recherche. Ensuite nous décrirons brièvement notre solution. Puis, nous décrirons l'expérience que nous avons menée et discuterons les résultats que nous avons observés. Et finalement, dans la conclusion nous rappellerons les points principaux de ce travail et suggérerons ce qu'un travail futur pourrait apporter.

6 Travaux relatifs

A notre connaissance, aucun travail actuel ne traite du sujet de cet article. Par contre, plusieurs articles abordent des sujets proches de celui-ci comme par exemple : [3], [4], [5] et [6].

7 Solution

Afin de vérifier si les critères capturent bien l'essence de l'esthétique, nous avons conduit une expérimentation sur un échantillon de 14 personnes selon la méthode vue au cours. Il leur a été demandé d'évaluer les critères pour ces interfaces et d'évaluer les interfaces subjectivement. Sur base de leurs réponses, nous avons fait une analyse statistiques qui nous a permis de tirer certaines conclusions.

8 Expérience

8.1 Méthode

Nous avons sélectionné 5 interfaces qui devraient être évaluées différemment par les critères et être intéressantes à analyser. Par exemple, a priori l'interface Ryanair devraient être mal évaluée par les critères tandis que celle de facebook devrait être bien évaluée.

Voici les interfaces analysées (voir les annexes pour des captures d'écran des interfaces) :

- Mur du profil facebook
- Page d'accueil Ryanair
- Trois interfaces du site Toyota Prius

Ensuite, nous avons composé un questionnaire reprenant, d'une part les 29 critères et d'autre part 6 ques-

tions subjectives (voir les annexes pour un exemple de questionnaire).

Nous avons sélectionné un échantillon de 14 personnes : 7 hommes et 7 femmes d'âge et d'origine différents.

Nous avons également questionné les utilisateurs pour établir leur profil dans le but de différencier les résultats en fonction des profils et de voir si le profil est déterminant dans la manière dont une interface est évaluée.

Pour les différents calculs, nous avons utilisé l'outil SPSS.

8.2 Résultats et discussions

8.2.1 Analyse de la corrélation entre les métriques et l'appréciation esthétique

Nous avons souhaité étudier la corrélation entre l'appréciation générale (esthétique) des interfaces et les notes données aux métriques. Pour cela nous avons utilisé une nouvelle variable calculée comme la somme des métriques pour chaque utilisateur. Avec SPSS nous avons calculé la covariance bivariée entre ces deux variables (la somme des métriques, et la moyenne des appréciations pour chaque interface). Nous affichons la corrélation de Pearson dans la Figure 1. Sur ce graphique nous y avons inclus les deux moyennes des variables étudiées. La moyenne des sommes des métriques a été divisée par le nombre de métriques (29) pour avoir des valeurs comprises entre 0 et 5.

On peut directement voir sur les Figure 1 et Figure 2 que la corrélation (en bleu) varie fort en fonction des interfaces. Elle est très bonne pour l'interface Prius Tracer ainsi que pour Facebook. Le résultat le plus étrange est celui de l'interface Prius Video. En effet c'est une interface avec une bonne moyenne d'appréciation (avec un score de 3,46/5 mais placée en 4ème position) mais qui ne semble pas du tout suivre les métriques. C'est une interface en effet assez atypique, qui ne semble pas suivre la logique des métriques.

Si l'on compare les écart-types et la corrélation de Pearson, on peut trouver une légère tendance à ce que la corrélation soit inversement proportionnelle aux écart-types. Plus il y a d'ambiguïté dans les réponses, moins la corrélation est importante.

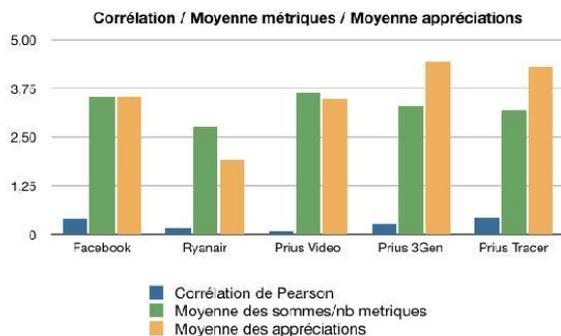


FIG. 1 – Analyse de la corrélation avec les moyennes

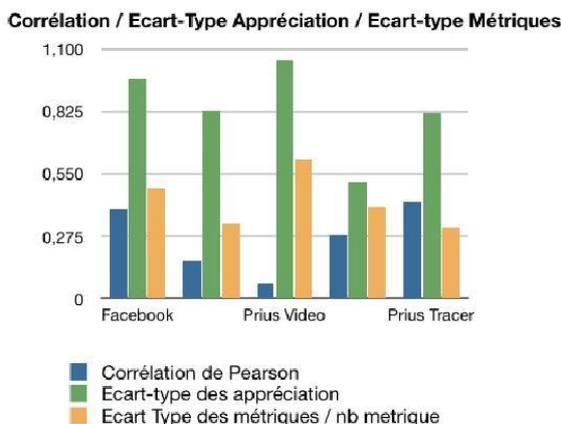


FIG. 2 – Analyse de la corrélation avec les écart-types

8.2.2 Indépendance des critères

Dans cette section, nous analysons l'indépendance des différents critères en calculant, pour quelques couples de critères, la corrélation de Pearson.

Pour cela, nous avons calculé les moyennes pour chaque critère et pour chaque interface en considérant l'échantillon total : soient 14 personnes plus les 5 membres de notre groupe. Après calcul des moyennes, nous avons donc 5 valeurs par critère, soit une valeur par interface.

Nous avons sélectionné quelques couples de variables qui sont censées être corrélées et d'autres qui ne sont pas censées l'être. Les valeurs des corrélations sont rapportées dans le Tableau 1.

Simplicité-Economie Une corrélation quasi parfaite est calculée. Ceci est en accord avec notre intuition. En

effet, le minimalisme/économie est souvent synonyme de simplicité en terme de communication. Pour une interface graphique cela s'applique également.

Economie-Sous-entendu Une corrélation significative est calculée. Ceci est en accord avec notre intuition. En effet, le sous-entendu et la suggestion sont souvent synonymes d'économie. Si l'interface suggère des choses sans devoir utiliser trop de widgets ou trop de texte, cela est souvent perçu comme de l'économie par l'utilisateur.

Simplicité-Cohérence Une corrélation quasi parfaite est calculée. Ceci est en accord avec notre intuition. En effet, si l'utilisateur juge l'interface cohérente, cela signifie qu'il la comprend. Il aura donc tendance à la trouver simple également.

Balance-Symétrie Une corrélation quasi parfaite est calculée. Ceci est en accord avec notre intuition. En effet, le critère de balance est une condition nécessaire pour avoir le critère de symétrie. Les deux critères sont donc corrélés par définition.

Cohérence-Neutralité Une corrélation tout de même significative est calculée. Nous pensions ces critères non-corrélés, le calcul nous indique le contraire. Toutefois, globalement, ces critères sont là pour qualifier la qualité d'une interface. Une bonne interface qui est globalement bonne va donc avoir tendance à satisfaire la plupart des critères ce qui peut expliquer qu'on observe une corrélation entre deux critères a priori décorés.

8.2.3 Classement des interfaces

Toujours, dans le but de savoir à quel point les critères capturent l'essence de l'esthétique nous avons comparé le classement subjectif des interfaces par les utilisateurs avec le classement établi à partir des données statistiques. Ces classements étant complètement différents, nous pouvons émettre des doutes quant à la façon dont le classement établi à partir des données statistiques a été effectué. En effet, ce classement a été fait en calculant une simple moyenne des critères. Une meilleure façon de procéder aurait été de calculer une moyenne pondérée en mettant des facteurs de poids fort pour les critères les plus significatifs.

Toutefois, selon la méthode utilisée, les critères ne captureraient pas forcément l'essence de l'esthétique et ne permettraient pas de prédire comment les utilisateurs perçoivent une interface. Nous présentons dans la Figure 3 une moyenne du classement subjectif des utilis-

teurs et dans la Figure 4 le classement établi à partir des évaluations des critères.

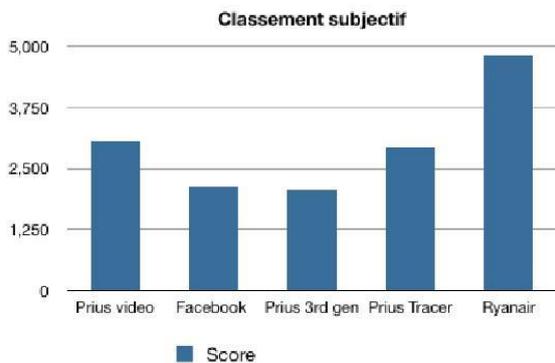


FIG. 3 – Classement subjectif établi par les utilisateurs

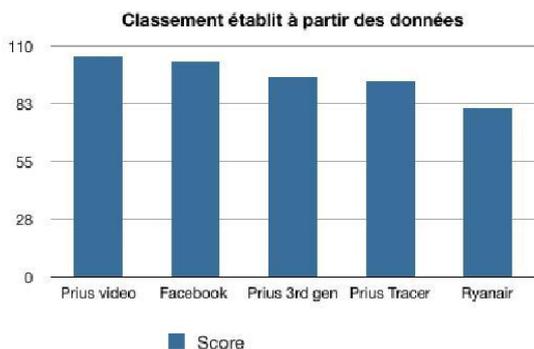


FIG. 4 – Classement établi à partir des données

Nous avons aussi fait une curieuse observation : c'est que le classement établi à partir de la question "trouvez-vous que l'interface est belle/esthétique?" était différent du classement obtenu en demandant aux utilisateurs de classer les interfaces dans l'ordre de préférence. Ceci semble suggérer une subjectivité extrême. Et nous fait douter quant à la mesurabilité de l'esthétique (voir Figure 5 pour avoir le classement établi à partir de la question "Trouvez-vous que l'interface est belle/esthétique?").

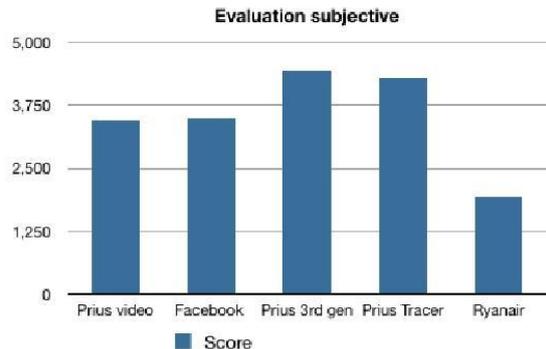


FIG. 5 – Réponse à la question "L'interface est-elle belle/esthétique?"

8.2.4 Facteur d'influence sur l'évaluation des critères d'esthétique

Pour cette partie du rapport nous avons voulu déterminer l'influence de certains facteurs qui caractérisent les personnes que nous avons interrogé sur l'évaluation des critères d'esthétiques.

Les différents facteurs que nous avons tenté de mettre en évidence sont les suivants :

- L'origine qui exprime le fait que la personne interrogée appartient ou non au continent européen.
- Le sexe qui détermine si la personne interrogée est un homme ou une femme.
- L'âge qui détermine si la personne interrogée a plus ou moins de 30 ans.

Pour déterminer l'influence de ces facteurs sur l'évaluation des critères d'esthétique, nous avons réalisé des tests d'hypothèse. Pour ce faire, nous avons pour chaque facteur et pour chaque interface partagé en deux ensembles les personnes interrogées afin de calculer la moyenne des sommes de tous les critères dans les deux ensembles pour ensuite déterminer H_0 , l'hypothèse statistique valant la différence de ces deux moyennes.

L'hypothèse nulle consistant à dire que la différence de ces deux moyennes vaut zéro correspondant à l'hypothèse où le facteur considéré n'a aucune influence sur l'évaluation des critères d'esthétique.

L'analyse de ces différents tests d'hypothèse nous permet d'établir le graphe présenté dans la Figure 6.

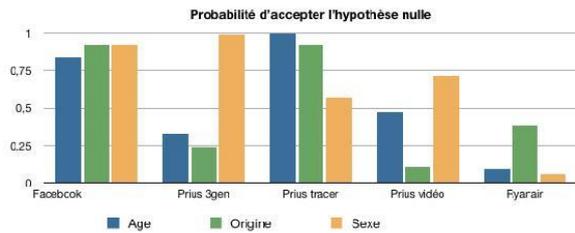


FIG. 6 –

Interprétation du graphe : ces valeurs correspondent à la probabilité d'accepter l'hypothèse nulle lorsque celle-ci est vérifiée.

Une critique que l'on pourrait faire sur nos tests d'hypothèse est que vu la petite taille de nos échantillons, nous ne pouvons pas tirer de conclusion très générale quant à l'évaluation des critères d'esthétique à grande échelle.

8.2.5 Esthétique et utilisabilité des interfaces

Pour évaluer la force du lien entre l'esthétique et l'utilisabilité perçue d'une interface, nous avons calculé la corrélation de Pearson de ces deux critères (évalués subjectivement par les utilisateurs). Une comparaison des différentes corrélations est présentée dans la Figure 7. La valeur de cette corrélation est donnée pour les différentes interfaces dans la Figure 6. Globalement, on peut conclure que la corrélation est moyennement élevée ce qui signifie qu'un utilisateur peut trouver subjectivement une interface non-esthétique mais facile à utiliser et inversement.

L'interface de Ryanair, par exemple, présente une faible corrélation entre utilisabilité et esthétique : elle est généralement mal évaluée en terme d'esthétique mais elle n'est pas forcément perçue comme non utilisable.

Bien sûr notre conclusion est forcément biaisée par la méthode que nous avons utilisé. Car dans un travail de plus grande envergure, on pourrait songer à une meilleure manière de mesurer l'utilisabilité perçue que simplement de demander "pensez-vous que cette interface est facilement utilisable?".

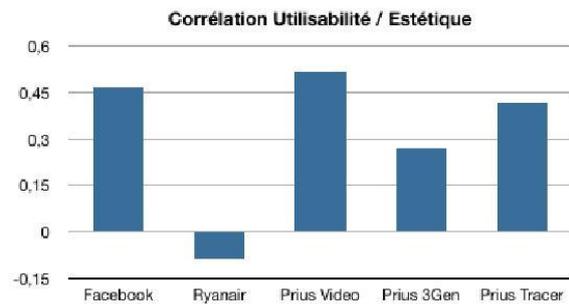


FIG. 7 – Corrélation utilisabilité-esthétique

	Corrélation de Pearson
Simplicité-Economie	0,955
Economie-Sous-entendu	0,870
Simplicité-Cohérence	0,918
Balance-Symétrie	0,901
Cohérence-Neutralité	0,600

TAB. 1 – Qualification de la redondance des critères

9 Conclusion

Notre travail met en évidence les observations suivantes :

- Une appréciation subjective positive de l'esthétique semble impliquer une évaluation plutôt positive des critères.
- Certains critères semblent être redondants. C'est-à-dire que ils ont tendance à être évalués de la même manière pour des mêmes interfaces.
- Notre méthode n'a pas permis de démontrer que les critères capturent bien la valeur esthétique d'une interface.
- Nous n'avons pas pu mettre en évidence un lien entre les origines, l'âge, le sexe et l'évaluation des critères par les utilisateurs.
- Globalement, une corrélation moyenne a été observée entre l'utilisabilité et l'esthétique d'une interface.

Travaux futurs :

- Faire un outil basé sur ces critères-ci et/ou d'autres qui mesure l'esthétique.
- Étudier plus en profondeur le lien entre utilisabilité et esthétique.
- Étudier plus en profondeur les "layout frames" (pour les interfaces multimedia).
- Refaire les expériences avec un grand échantillon.

10 Remerciements

Nous tenions à adresser nos remerciements aux interviewés pour leur disponibilité ainsi qu'au professeur J. Vanderdonckt pour son assistance.

Références

1. J. Hartmann, A. Sutcliffe et A. de Angeli, Towards a Theory of User Judgment of Aesthetics and User Interface Quality - 2008
2. J. Vanderdonckt, X. Gillo, Visual Techniques for Traditional and Multimedia Layouts - 1994
3. D. Check Ling Ngo, L. Seng Teo et J. G. Byrne, A Mathematical Theory of Interface Aesthetics
4. D. Check Ling Ngo et J. G. Byrne, Another Look at a Model for Evaluating Interface Aesthetics - 2001
5. R. Mahajan et B. Shneiderman, A Family of User Interface Consistency Checking Tools - 1995
6. N. Tractinsky, A.S. Katz, D. Ikar, What is Beautiful is Usable - 2000

Annexes

Évaluation générale

* Required

Données personnelles

Nom *

Prenom *

Évaluation

Faites un classement des 5 interfaces du plus esthétique au moins esthétique

Facebook *

- 1ère place
- 2de place
- 3ème place
- 4ème place
- 5ème place

Ryanair *

- 1ère place
- 2de place
- 3ème place
- 4ème place
- 5ème place

Prius 3rd Gen *

- 1ère place
- 2de place
- 3ème place
- 4ème place
- 5ème place

FIG. 8 – Exemple de questionnaire utilisé pour interroger les utilisateurs

Evaluation de l'interface

L'interface est-elle belle/esthétique? *

1 2 3 4 5

Pas belle Belle

Citer trois choses que vous avez aimé de cette interface *

Citer trois choses que vous n'avez pas aimé de cette interface *

L'interface est-elle originale ou standard? *

1 2 3 4 5

Standard Originale

L'interface "fait elle professionnel" ou amateur *

1 2 3 4 5

Amateur Professionnel

Pensez vous que le site est facilement utilisable? *

1 2 3 4 5

Difficilement utilisable Facilement utilisable

FIG. 9 – Evaluation subjective des interfaces

Données personnelles

Nom *

Prénom *

Origine ethnique *

- Asie
- Afrique du nord
- Afrique subsaharienne
- Amérique latine
- Canada et les États-Unis
- Europe
- Moyen-Orient
- Océanie
- Other:

Éducation *
Niveau d'instruction

- Primaire
- Secondaire
- Technique
- Universitaire
- Other:

Évaluation de l'interface

L'interface est-elle belle/esthétique? *

1 2 3 4 5

Pas belle Belle

FIG. 10 – Etablissement du profil des utilisateurs

Principes graphiques de la conception visuelle

FACEBOOK

Données personnelles

Nom

Prénom

Age

Sexe
 Homme Femme

Techniques physiques

Balancé
1 2 3 4 5
Peu Beaucoup

Symétric
1 2 3 4 5
Peu Beaucoup

Régularité
1 2 3 4 5
Peu Beaucoup

Alignement
1 2 3 4 5
Peu Beaucoup

Proportion
1 2 3 4 5
Peu Beaucoup

criteres.png

FIG. 11 – Questionnaire d'évaluation des critères

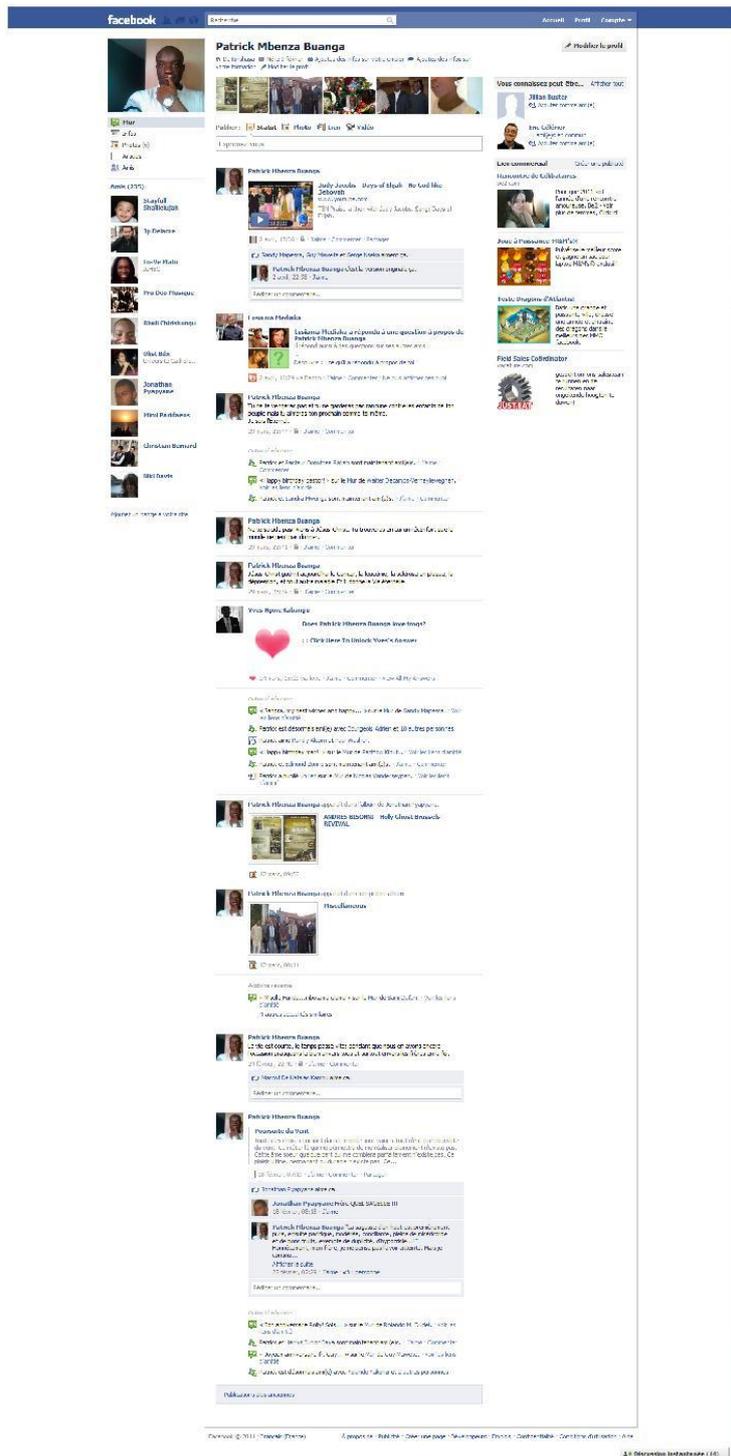


FIG. 12 – Mur du profil facebook

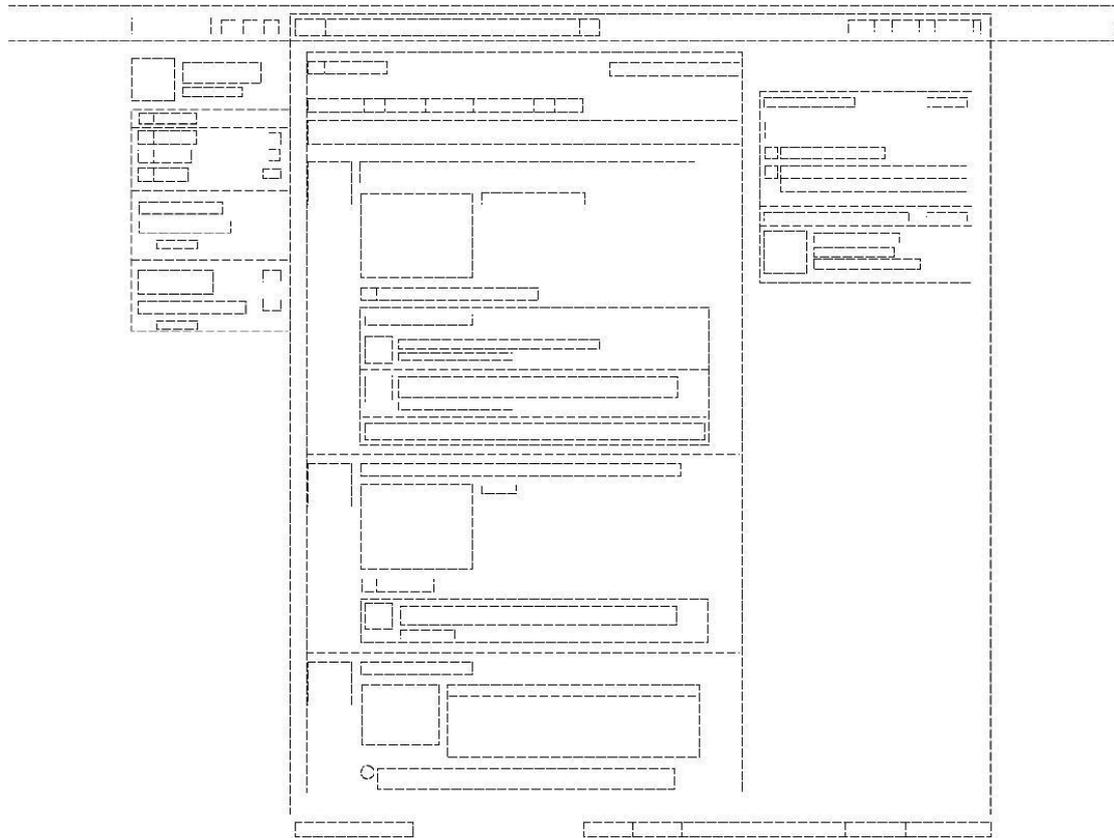


FIG. 13 – Grillage de l'interface du mur du profil facebook

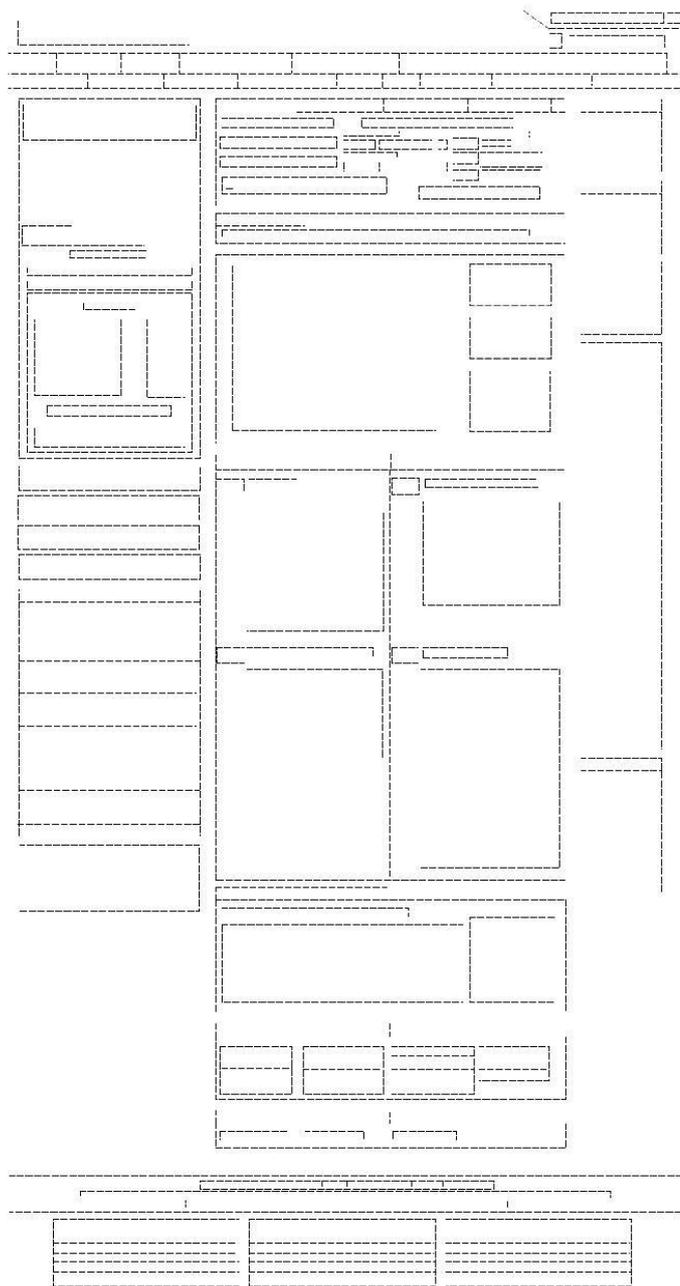


FIG. 15 – Grillage Accueil Ryanair

3rd Generation
prius

Touch Tracer Display
Control everything inside. Miss nothing outside.

[Feature Video](#)

Movements on the touchpads on the Touch Tracer Display are visible at eye level to help keep the driver's attention on the road.

Aerodynamic Styling Solar Roof	Interior Refinements Touch Tracer Display	3rd Gen Improvements EV/Eco/Power Modes	Advanced Parking Guidance System Lane Keep Assist	Pre-Collision System Safety Features
EXTERIOR	INTERIOR	MPG INNOVATION	TECHNOLOGY	SAFETY

© 2009 Toyota Motor Sales, U.S.A., Inc. All information presented herein applies to U.S. vehicles only. Vehicle shown with optional equipment. Production model may vary. An ideal environment for the scale.

TOYOTA
moving forward

[YOUR PRIVACY RIGHTS](#) | [CLICK FOR DETAILS](#) | [VISIT TOYOTA.COM/PRIUS](#) | [BUILD & PRICE](#)

FIG. 16 – Prius : Touch Tracer

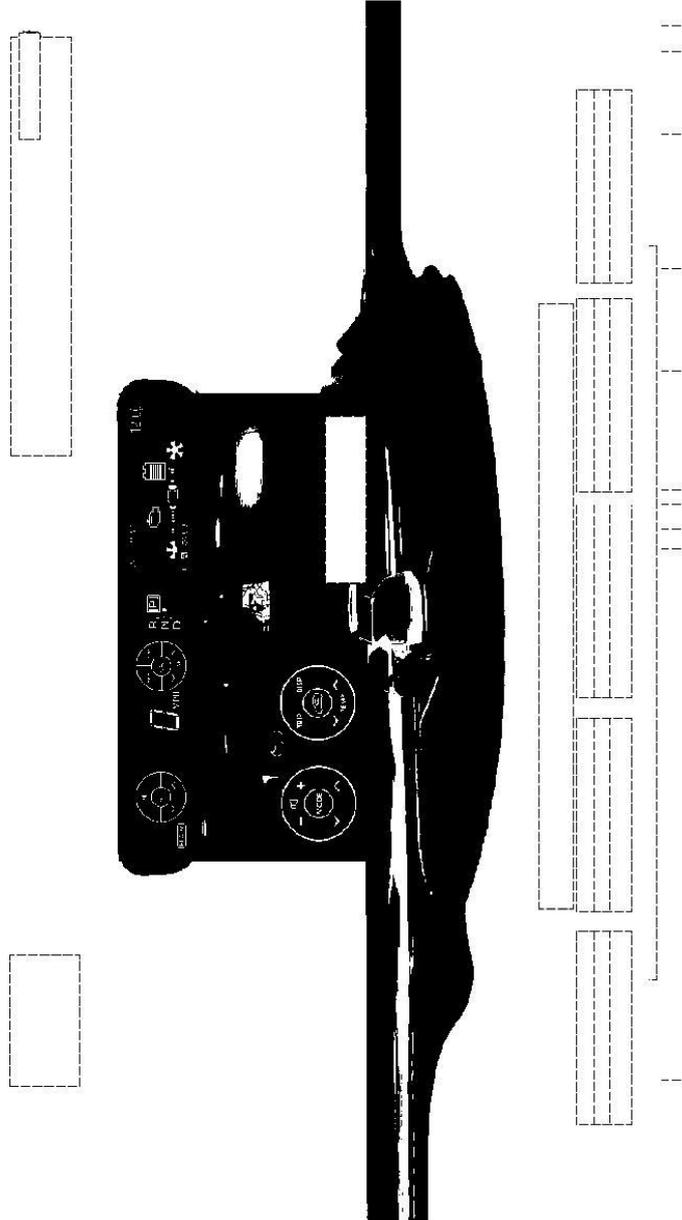


FIG. 17 – Prius : Touch Tracer

3rd Generation prius

[Feature Video 1](#)
[Feature Video 2](#)

3rd Generation Improvements
 How do you receive an EPA 50 mpg rating*?

*2013 EPA 50 mpg city/45 mpg highway/50 mpg combined. Actual mileage will vary.

Adjust the sliders to see what makes the 50 mpg rating possible.

ENGINE

ELECTRIC MOTOR

BATTERY

+
-

From starting of thorough light acceleration, full hybrids like Prius can be powered solely by their electric motor, as opposed to major hybrids which require that their engines turn while driving.

Aerodynamic Styling	Interior Refinements	EV/Eco/Power Modes	3rd Gen Improvements
Solar Roof	Touch Tracer Display	Lane Keep Assist	Advanced Parking Guidance System
EXTERIOR	INTERIOR	MFG INNOVATION	TECHNOLOGY
			Pic-Collision System
			Safety Features
			SAFETY

© 2009-2017 Toyota Motor Sales, U.S.A., Inc. All information presented herein applies to U.S. vehicles only. Vehicle shown with optional equipment. Production model may vary. Altered environment for scale.

strong forward

[YOUR PRIVACY RIGHTS](#)
[CLICK FOR DETAILS](#)

[VISIT TOYOTA.COM/PRIUS](#)
[BUILD & PRICE](#)

FIG. 18 – Prius 3rd generation improvement

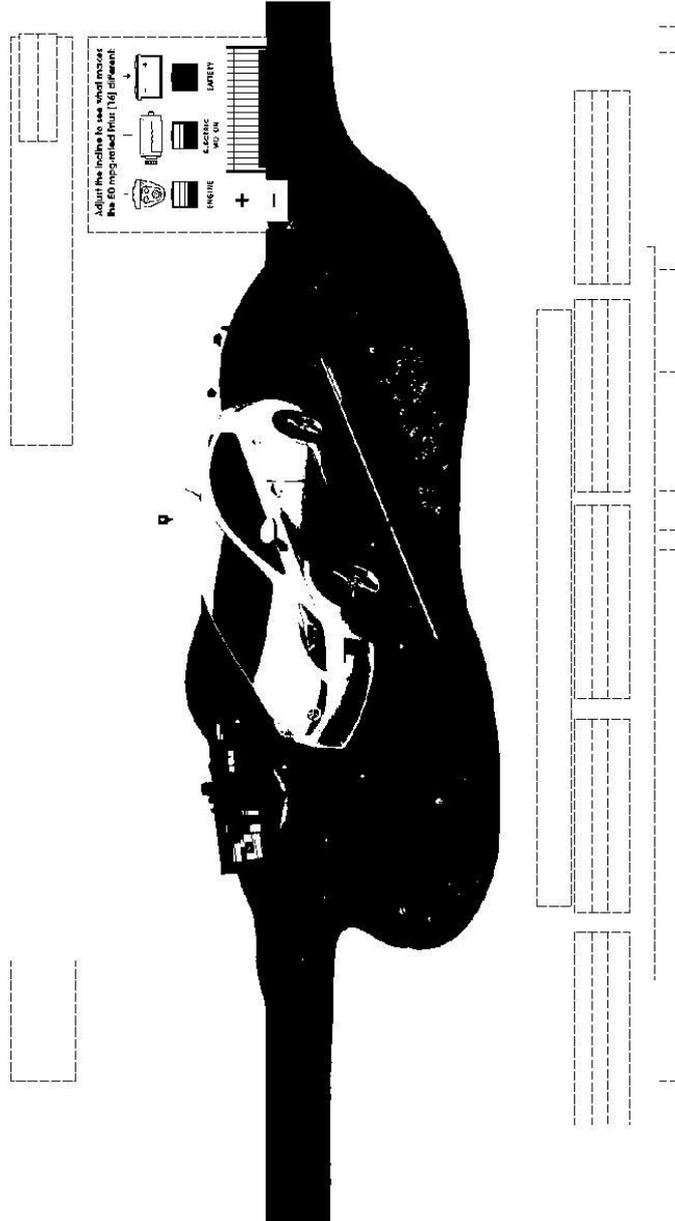


FIG. 19 – Prius 3rd generation improvement

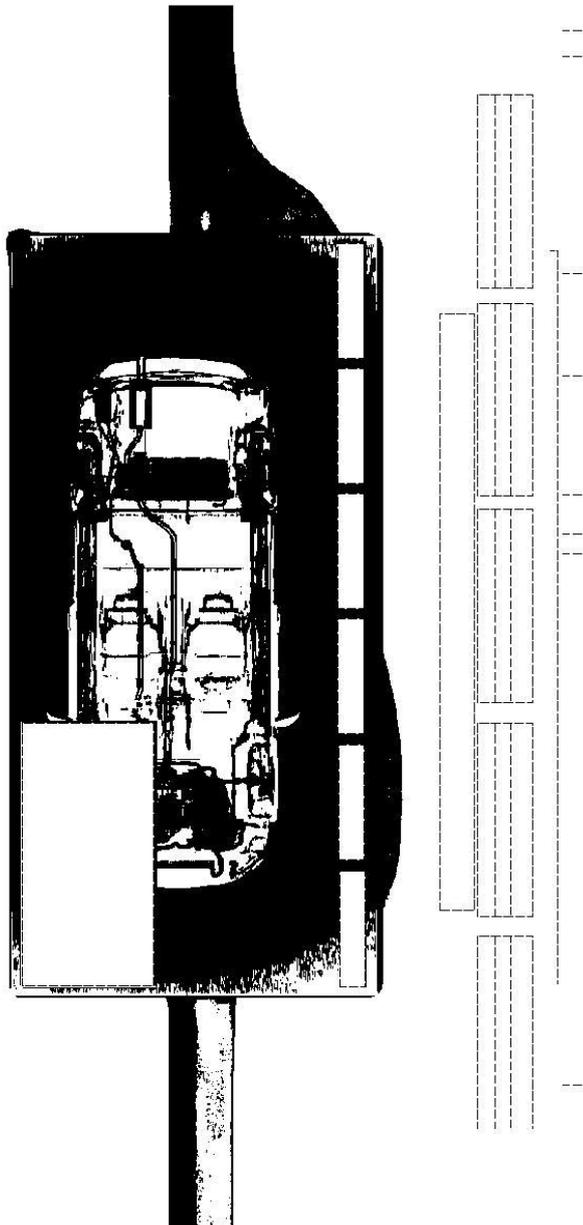


FIG. 21 - Prius 3rd generation : video

Appendix C : Selected Metrics.

1. Unity

Aesthetics Metric

❖ Metric Name

Unity

❖ References

[2] D.C.L. Ngo, L.S. Teo, J. G. Byrne *A Mathematical Theory of Interface Aesthetics* available at <http://www.mi.sanu.ac.rs/vismath/ngo/index.html>

[16] Vanderdonckt, J. "Visual Design Methods in Interactive Applications", 2002

DEFINITION

Unity is coherence, a totality of elements that is visually all one piece. With unity, the elements seem to belong together, to dovetail so completely that they are seen as one thing. Unity is achieved by using similar sizes, shapes, or colors for related information and leaving less space between elements of a screen than the space left at the margins. Unity, by definition, is the extent to which the screen elements seem to belong together.[2]

Unity is the placement of individual IO into one totality (e.g. a window) that is visually all of a piece (fig.17a). With unity, all IO seem to belong to each other and to be bound so that they can be seen as a whole and taken as one sealed unsectile thing: seeing one element is seeing the whole. Unity can be revealed with similar sizes, colors, surrounding blank spaces, logical organization exhibiting interrelation of IO in terms of the whole.[16]

- *Comments: Grouping* [16] contributes to unity; but Unity can be achieved without grouping.
- *Importance/Weight: medium*

❖ Formula

$$UM = \frac{|UM_{form}| + |UM_{space}|}{2} \in [0,1]$$

UM_{form} is the extent to which the objects are related in size, shape, and colour with

$$UM_{form} = 1 - \frac{n_{size} + n_{colour} + n_{shape} - 3}{3n}$$

and UM_{space} is a relative measure of the space between groups and that of the margins with

$$UM_{space} = 1 - \frac{a_{layout} - \sum_i^n a_i}{a_{frame} - \sum_i^n a_i}$$

where a_i , a_{layout} , and a_{frame} are the areas of object i , the layout, and the frame, respectively; n_{size} , n_{colour} , and n_{shape} are the numbers of sizes, colours, and shapes used, respectively; and n is the number of objects on the frame.

❖ Motivation

All the widgets in the interface below perfectly fit together and seem to belong to a whole entity. This unity is aesthetically pleasing in the author's view.

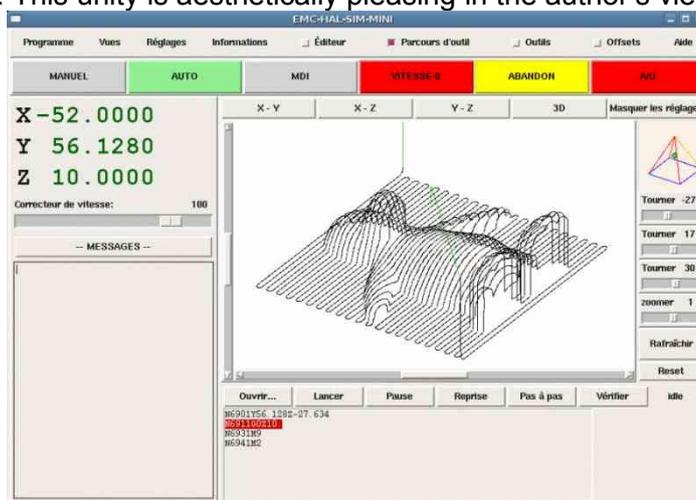


Figure 39: Perfect unity.

The interface in fig.26 is unaesthetic for many reasons. One of them is probably the lack of unity. The images/links on the lower half of the screen do not fit together as smoothly as in the interface above.



Figure 40: Poor unity.

❖ Advantages

- ...

❖ Drawbacks

- According to the formula, better UM_{form} values are associated with fewer object sizes used; UM_{form} is 1 if the layout uses only one size. Better UM_{space} values are closely related to how tightly the screen is packed; UM_{space} is 1 when there is no space between objects and it is 0 when there is no space left at the margins... this is quite an (over)simplified way of measuring unity!

2. Symmetry

Aesthetics Metric

❖ Metric Name

Symmetry

❖ References

[2] D.C.L. Ngo, L.S. Teo, J. G. Byrne *A Mathematical Theory of Interface Aesthetics* available at <http://www.mi.sanu.ac.rs/vismath/ngo/index.html>
 [16] Vanderdonckt, J. "Visual Design Methods in Interactive Applications", 2002

DEFINITION

- Symmetry is axial duplication: A unit on one side of the centre line is exactly replicated on the other side. Vertical symmetry refers to the balanced arrangement of equivalent elements about a vertical axis, and horizontal symmetry about a horizontal axis. Radial symmetry consists of equivalent elements balanced about two or more axes that intersect at a central point.[2]

Symmetry consists of duplicating the visual image of IO along a horizontal (fig. 27a) and/or vertical axis (fig. 27b) (e.g. left on the right, top to bottom, or vice versa).

Achieving symmetry automatically preserves balance, but the balance can be performed without symmetry. Symmetry (fig. 28a) is very simple to verify and logical to imagine, but can lead to static layouts without originality. The opposite of symmetry is asymmetry (fig. 28b) where at least one IO does not possess a replication on the other axis side.

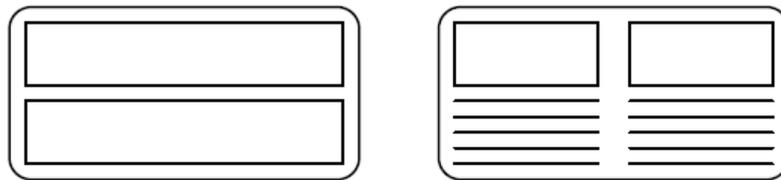


Figure 41: Horizontally and vertically symmetric layouts.

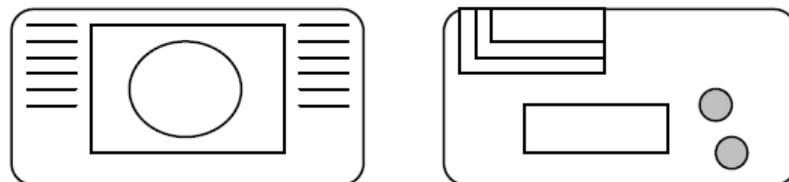


Figure 42: Symmetric and asymmetric layouts.

- *Comments:*
 - o Achieving symmetry automatically preserves balance, but balance can be performed without symmetry.
 - o This interface is stable but not symmetric. So stability can be achieved without symmetry.



Figure 43: Asymmetric but stable layout.

- Importance/Weight: low

❖ Formula

Symmetry is given by

$$SYM = 1 - \frac{|SYM_{vertical}| + |SYM_{horizontal}| + |SYM_{radial}|}{3} \in [0,1] \quad (9)$$

$SYM_{vertical}$, $SYM_{horizontal}$, and SYM_{radial} are, respectively, the vertical, horizontal, and radial symmetries with

$$SYM_{vertical} = \frac{|X'_{UL} - X'_{UR}| + |X'_{LL} - X'_{LR}| + |Y'_{UL} - Y'_{UR}| + |Y'_{LL} - Y'_{LR}| + |H'_{UL} - H'_{UR}| + |H'_{LL} - H'_{LR}| + |B'_{UL} - B'_{UR}| + |B'_{LL} - B'_{LR}| + |\ominus'_{UL} - \ominus'_{UR}| + |\ominus'_{LL} - \ominus'_{LR}| + |R'_{UL} - R'_{UR}| + |R'_{LL} - R'_{LR}|}{12} \quad (10)$$

$$SYM_{horizontal} = \frac{|X'_{UL} - X'_{LL}| + |X'_{UR} - X'_{LR}| + |Y'_{UL} - Y'_{LL}| + |Y'_{UR} - Y'_{LR}| + |H'_{UL} - H'_{LL}| + |H'_{UR} - H'_{LR}| + |B'_{UL} - B'_{LL}| + |B'_{UR} - B'_{LR}| + |\ominus'_{UL} - \ominus'_{LL}| + |\ominus'_{UR} - \ominus'_{LR}| + |R'_{UL} - R'_{LL}| + |R'_{UR} - R'_{LR}|}{12} \quad (11)$$

$$SYM_{radial} = \frac{|X'_{UL} - X'_{LR}| + |X'_{UR} - X'_{LL}| + |Y'_{UL} - Y'_{LR}| + |Y'_{UR} - Y'_{LL}| + |H'_{UL} - H'_{LR}| + |H'_{UR} - H'_{LL}| + |B'_{UL} - B'_{LR}| + |B'_{UR} - B'_{LL}| + |\ominus'_{UL} - \ominus'_{LR}| + |\ominus'_{UR} - \ominus'_{LL}| + |R'_{UL} - R'_{LR}| + |R'_{UR} - R'_{LL}|}{12} \quad (12)$$

where X'_i , Y'_i , H'_i , B'_i , Q'_i , and R'_i are, respectively, the normalised values of X_j , Y_j , H_j , B_j , Q_j , and R_j with

$$X_j = \sum_i^{n_j} |x_{ij} - x_c| \quad j = UL, UR, LL, LR \quad (13)$$

$$Y_j = \sum_i^{n_j} |y_{ij} - y_c| \quad j = UL, UR, LL, LR \quad (14)$$

$$H_j = \sum_i^{n_j} h_{ij} \quad j = UL, UR, LL, LR \quad (15)$$

$$B_j = \sum_i^{n_j} b_{ij} \quad j = UL, UR, LL, LR \quad (16)$$

$$\Theta_j = \sum_i^{n_j} \left| \frac{y_{ij} - y_c}{x_{ij} - x_c} \right| \quad j = UL, UR, LL, LR \quad (17)$$

$$R_j = \sum_i^{n_j} \sqrt{(x_{ij} - x_c)^2 + (y_{ij} - y_c)^2} \quad j = UL, UR, LL, LR \quad (18)$$

where *UL*, *UR*, *LL*, and *LR* stand for upper-left, upper-right, lower-left, and lower-right, respectively; (x_{ij}, y_{ij}) and (x_c, y_c) are the co-ordinates of the centres of object *i* on quadrant *j* and the frame; b_{ij} and h_{ij} are the width and height of the object; and n_j is the total number of objects on the quadrant.

❖ Motivation

This interface is unstable because it's not symmetric:



Figure 44: Unstable interface.

The same interface looks better when made symmetric:



Figure 45: Same interface; edited in Microsoft Paint.

❖ Advantages

- Symmetry implies balance.

❖ Drawbacks

- The formula is too complex.
- Too much symmetry is bad because it leads to rigidity:

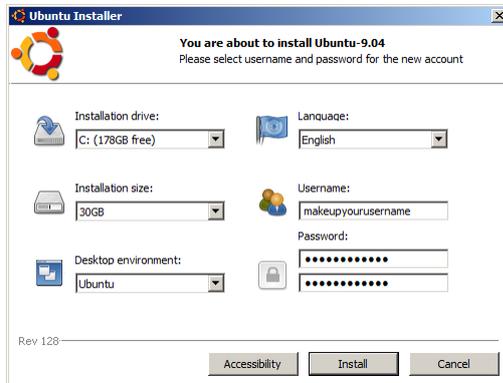


Figure 46: This interface is too symmetric!

3. Alignment

Aesthetics Metric

❖ Metric Name

Alignment

❖ References

[16] Vanderdonckt, J. "Visual Design Methods in Interactive Applications", 2002

DEFINITION

Alignment is guaranteed if the number of vertical alignment points in a row and the number of horizontal alignment points in a column is reduced, minimized (fig. 7).

Misalignment - the opposite of the alignment - occurs when the number of alignment points is significantly high (fig. 7).

Misalignment is accentuated when all IO containing task's data are placed just after their identification labels. Fonts with descenders and ascenders may affect alignment of similar IO if badly used. [16]

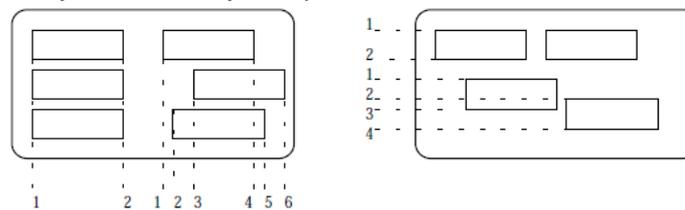


Fig. 7. Vertical, horizontal alignments and misalignments.

- *Comments:* none.
- *Importance/Weight:* high

❖ Formula

There's no formula provided in the literature. We had to make one ourselves:

$$4n - (nhap * nvap / n) / 4n - 4$$

❖ Motivation

We wondered why the interface below was so unaesthetic. After just making one of the lines of the inner rectangle straight using Microsoft Paint, the interface looked much more attractive. We were surprised to see how alignment could so dramatically change the aesthetic quality of a design.

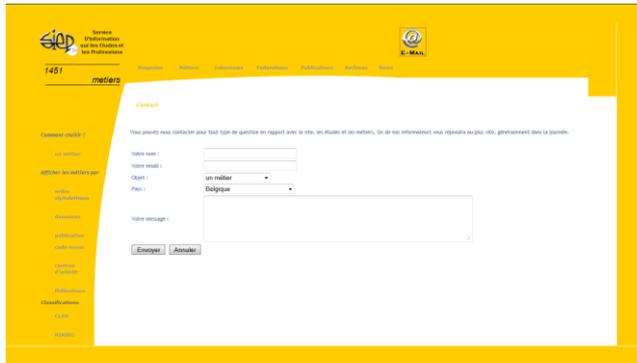


Figure 47: Unaesthetic because of bad alignment.



Figure 48: Same interface aligned with Microsoft Paint.

❖ Advantages

- Alignment is a very accessible metric (easy to understand and implement) But a very powerful one. An aligned interface looks neat and clean.

❖ Drawbacks

- The formula has been refined several times, but we have not been able to find a perfect metric to detect misaligned widgets.

4. Grouping

Aesthetics Metric

❖ Metric Name

Grouping

❖ References

[16] Vanderdonckt, J. "Visual Design Methods in Interactive Applications", 2002

DEFINITION

- Grouping is a visual technique that creates a circumstance of give and take of relative interaction (fig. 35a). Grouping is mainly based on the law of attraction: two grouped IO fight for attention in their interaction by establishing individual statements depending on the distance between the IO. The closer the IO are, the stronger the attraction is. Grouping is also affected by the law of similarity: opposite, dissimilar IO repel each other, but equal, similar IO attract each other. When dissimilar IO are grouped, the human eye increases the relation between them. When similar and dissimilar IO are grouped, hidden connections are identified as fast. Grouping is one of the best techniques for structuring a layout namely by providing an aesthetic appearance, by helping remembering and by accelerating a layout search.[16]

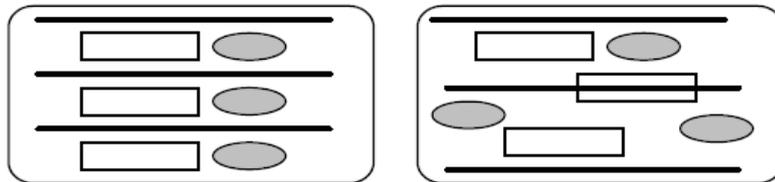


Figure 49: Grouped and split layouts.

- *Comments:* Grouping fosters unity. But unity can be achieved without grouping.
- *Importance/Weight:* medium

❖ Formula

There's no formula provided in the literature. We had to make one ourselves:
 $(0.85f) * UMspace + (0.15f) * AL$.

Where $UMspace = 1 - (aLayout - sumOfAreas / aFrame - sumOfAreas)$

Where $AL = \min(1, (4 * Ni - (nHap * nVap / Ni)) / (4 * Ni - 4))$

Where $aLayout$ is the area of the layout; $sumOfAreas$ is the sum of the areas of objects; $aFrame$ is the area of the frame; Ni is the number of objects of type i ; $nHap$ and $nVap$ are the number of horizontal alignment points and the number of vertical alignment points, respectively.

❖ Motivation

This interface convinced the author that grouping was an important characteristic of an interface; it has an excellent grouping.



Figure 50: Excellent grouping.

❖ Advantages

- ...

❖ Drawbacks

- The formula has been refined several times, but we have not been able to find a perfect metric to measure good grouping of widgets.

5. Homogeneity (Repartition)

Aesthetics Metric

❖ Metric Name

Homogeneity/Repartition

❖ References

[2] D.C.L. Ngo, L.S. Teo, J. G. Byrne *A Mathematical Theory of Interface Aesthetics* available at <http://www.mi.sanu.ac.rs/vismath/ngo/index.html>
[16] Vanderdonckt, J. "Visual Design Methods in Interactive Applications", 2002

DEFINITION

Homogeneity, by definition, is a measure of how evenly the objects are distributed among the quadrants.[2]

Repartition proposes to share IO among the four quadrants of the layout as systematically as possible (fig.19a). Experimental results argued it is not often the case while first, second, third, and fourth quadrants consume 40%, 20%, 15%, and 25% of the IO, respectively (fig. 18).[16]

40%	20%
25%	15%

Figure 51: IO Repartition among the screen quadrants.

- *Comments:* none
- *Importance/Weight:* low

❖ Formula

Homogeneity is given by

$$HM = \frac{W}{W_{\max}} \in [0,1]$$

W is the number of different ways a group of n objects can be arranged for the four quadrants when

n_j is the total number of objects in quadrant j , that is

$$W = \frac{n!}{\prod_{j=UL,UR,LL,LR} n_j} = \frac{n!}{n_{UL}! n_{UR}! n_{LL}! n_{LR}!}$$

W is maximum when the n objects are evenly allocated to the various quadrants of the screen, as

compared to more or less uneven allocations among the quadrants, and thus

$$W_{\max} = \frac{n!}{\frac{n}{4}! \frac{n}{4}! \frac{n}{4}! \frac{n}{4}!} = \frac{n!}{\left(\frac{n}{4}\right)^4}$$

where n_{UL} , n_{UR} , n_{LL} , and n_{LR} are the numbers of objects on the upper-left, upper-right, lower-left,

and lower-right quadrants, respectively; and n is the number of objects on the frame.

Where a_i and a_{frame} are the areas of object i and the frame; and n is the number of objects on the frame (or interface).

❖ Motivation

Equal repartition of objects among the four quadrants seemed to be obviously important to the author. But, after studying the interface below, It seems that the upper and lower left quadrants are more important than the quadrants on the right. So we modified the formula a little bit to give more importance to the upper left and lower left quadrants :



Figure 52: Quadrant preference.

❖ Advantages

- ...

❖ Drawbacks

- The formula for Homogeneity (Repartition) is based on the number of widgets in each quadrant of the screen; our tests showed that the metric would be more accurate if it was based on the area of objects instead of the number of objects.

6. Simplicity

Aesthetics Metric

❖ Metric Name

Simplicity

❖ References

[2] D.C.L. Ngo, L.S. Teo, J. G. Byrne *A Mathematical Theory of Interface Aesthetics* available at <http://www.mi.sanu.ac.rs/vismath/ngo/index.html>

[16] Vanderdonckt, J. "Visual Design Methods in Interactive Applications", 2002

DEFINITION

Simplicity is directness and singleness of form, a combination of elements that results in ease in comprehending the meaning of a pattern. Simplicity in screen design is achieved by optimizing the number of elements on a screen and minimizing the alignment points. [2]

Simplicity is directness and singleness of layout, free from secondary complications or sophistications (fig. 39a). Simplicity improves largely the ease of understanding the layout grid or frame. Simplicity is guaranteed by placing IO according to a logical and natural arrangement (e.g. by frequency, by physical property) driven by the task's semantics. Complexity - the opposite of simplicity - increases visual intricacy with too much units, forces and results and hinders any organisation of the layout grid (fig. 39b). [16]

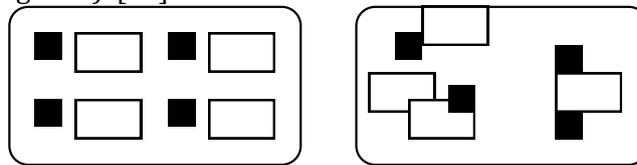


Figure 53: Simple and complex layouts.

- *Comments:* none
- *Importance/Weight:* medium

❖ **Formula**

An easier method of calculation is

$$SMM = \frac{3}{n_{vap} + n_{hap} + n} \in [0,1] \quad (43)$$

where n_{vap} and n_{hap} are the numbers of vertical and horizontal alignment points; and n is the number of objects on the frame.

❖ **Motivation**

The simplicity of this interface is aesthetically attractive in itself.



Figure 54: Simple interface.

❖ Advantages

- ...

❖ Drawbacks

- The formula provided by Ngo et al.[2] is way too simple. The authors admit that their formulas are not perfect and that they are constantly being studied and improved. The formula for simplicity needs much improvement according to us. Their formula for Simplicity is inversely proportional to the sum of the numbers of alignment points and screen objects: as the summation decreases, simplicity tends to increase. Higher values are associated with a smaller number of objects used and related to how well the objects are aligned. Simplicity is 1 if there is only one object used. However it's not hard to imagine a perfectly simple layout that contains more than one object.

7. Economy

Aesthetics Metric

❖ Metric Name

Economy

❖ References

[2]D.C.L. Ngo, L.S. Teo, J. G. Byrne *A Mathematical Theory of Interface Aesthetics* available at <http://www.mi.sanu.ac.rs/vismath/ngo/index.html>
 [16] Vanderdonckt, J. "Visual Design Methods in Interactive Applications", 2002

DEFINITION

Economy is the careful and discreet use of display elements to get the message across as simple as possible. Economy is achieved by using as few styles, displays techniques, and colors as possible.[2]

Economy is the frugal and judicious use of IO in the layout to present information as simply as possible. Economy can be pursued when necessary and sufficient IO are placed in the layout and nothing else : no IO that are extraneous to the user's task.[16]

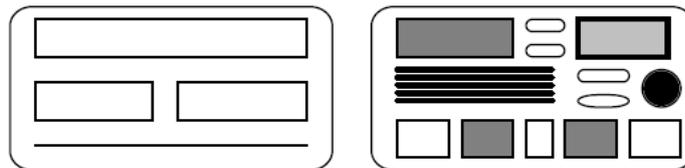


Figure 55: Uncluttered and cluttered layouts.

- *Comments:* none
- *Importance/Weight:* low

❖ Formula

$$ECM = \frac{3}{n_{size} + n_{colour} + n_{shape}} \in [0,1]$$

where n_{size} , n_{colour} , and n_{shape} are the numbers of sizes, colours, and shapes used, respectively.

❖ Motivation

This interface uses widgets and labels judiciously and economically. Only the minimum amount of objects and styles needed to convey the information is placed on the interface.

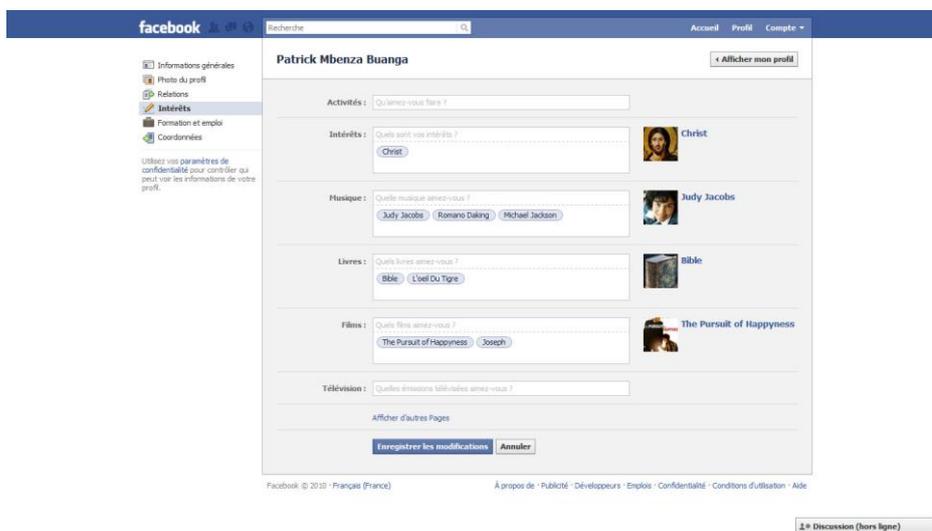


Figure 56: Excellent economy.

❖ Advantages

- The formula is simple enough and straightforward to implement.

❖ Drawbacks

- The formula provided by Ngo et al.[2] is way too simple. The authors admit that their formulas are not perfect and that they are constantly being studied and improved. The formula for economy needs much improvement according to us. If there's only one shape, color and size the result is 1 (perfect economy); the more varied the worse the result. However a varied set of components can be arranged in a very economic fashion.

8. Proportion

Aesthetics Metric

❖ Metric Name

Proportion

❖ References

- [2]D.C.L. Ngo, L.S. Teo, J. G. Byrne *A Mathematical Theory of Interface Aesthetics* available at <http://www.mi.sanu.ac.rs/vismath/ngo/index.html>
- [16] Vanderdonckt, J. "Visual Design Methods in Interactive Applications", 2002

DEFINITION

What constitutes beauty in one culture is not necessarily considered the same by another culture, but some proportional shapes have stood the test of time and are found in abundance today. Marcus describes the following shapes as aesthetically pleasing. [2]

- Square (1:1)
- Square root of two (1:1.414)
- Golden rectangle (1:1.618)
- Square root of three (1:1.732)
- Double square (1:2)

Proportion strives for an aesthetically appealing ratio between the dimensions of IO (often composite IO) (fig. 43a). Because dimensions exist in the real world, we can feel them, we can see them, we can compare them. Mirroring dimensions in layouts consists of retracing

this feeling, this illusion in the user interface. The ratio is calculated by dividing the height of an IO by its length.

Several proportions have been either proved aesthetic (e.g. the Golden Ratio) or widely

and conventionally preferred (e.g. 1:2, 1:1.29, 1:1.5, 1:4/3, 1:1.6). [16]

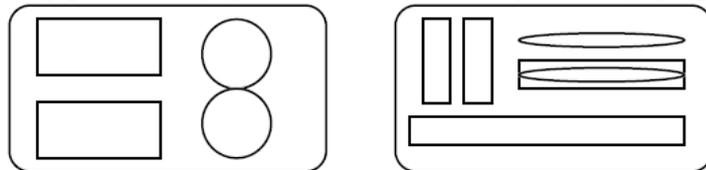


Figure 57: Proportioned and disproportioned layouts.

- *Comments:* none
- *Importance/Weight:* low

❖ Formula

Proportion, by definition, is the comparative relationship between the dimensions of the screen components and proportional shapes and is given by

$$PM = \frac{|PM_{object}| + |PM_{layout}|}{2} \in [0,1]$$

PM_{object} is the difference between the proportions of the objects and the closest proportional shapes described by Marcus with

$$PM_{object} = \frac{1}{n} \sum_i^n \left(1 - \frac{\min(|p_j - p_i|, j = sq, r2, gr, r3, ds)}{0.5} \right)$$

and PM_{layout} is the difference between the proportions of the layout and the closest proportional shape with

$$PM_{layout} = 1 - \frac{\min(|p_j - p_{layout}|, j = sq, r2, gr, r3, ds)}{0.5}$$

with

$$\{p_{sq}, p_{r2}, p_{gr}, p_{r3}, p_{ds}\} = \left\{ \frac{1}{1}, \frac{1}{1.414}, \frac{1}{1.618}, \frac{1}{1.732}, \frac{1}{2} \right\}$$

$$p_i = \begin{cases} r_i & \text{if } r_i \leq 1 \\ \frac{1}{r_i} & \text{Otherwise} \end{cases}$$

$$P_{layout} = \begin{cases} r_{layout} & \text{if } r \leq 1 \\ \frac{1}{r_{layout}} & \text{Otherwise} \end{cases}$$

with

$$r_i = \frac{h_i}{b_i}$$

$$r_{layout} = \frac{h_{layout}}{b_{layout}}$$

where b_i and h_i are the width and height of object i ; b_{layout} and h_{layout} are the width and height of the layout; and p_j is the proportion of shape j . (Note that the maximum value of $(p_j \cdot r_{layout})$ is 0.5.)

❖ Motivation

The definition says it all: some proportional shapes have stood the test of time and are found in abundance today.

❖ Advantages

- Simple metric; easy to understand and implement.

❖ Drawbacks

- ...

Appendix D : Form used to collect users ratings.

Interview Interfaces

Cette interface est-elle belle?

Android

- Belle
- Moyenne
- Laide

Printer 1

- Belle
- Moyenne
- Laide

Printer 2

- Belle
- Moyenne
- Laide

Guf 1

- Belle
- Moyenne
- Laide

Guf 2

- Belle
- Moyenne
- Laide

Classement

Android est ...

la plus belle ▼

Printer 1 est ...

la plus belle ▼

Printer 2 est ...

la plus belle ▼

Guf 1 est ...

la plus belle ▼

Guf 2 est ...

la plus belle ▼

Nom

Prénom

Age

Sexe

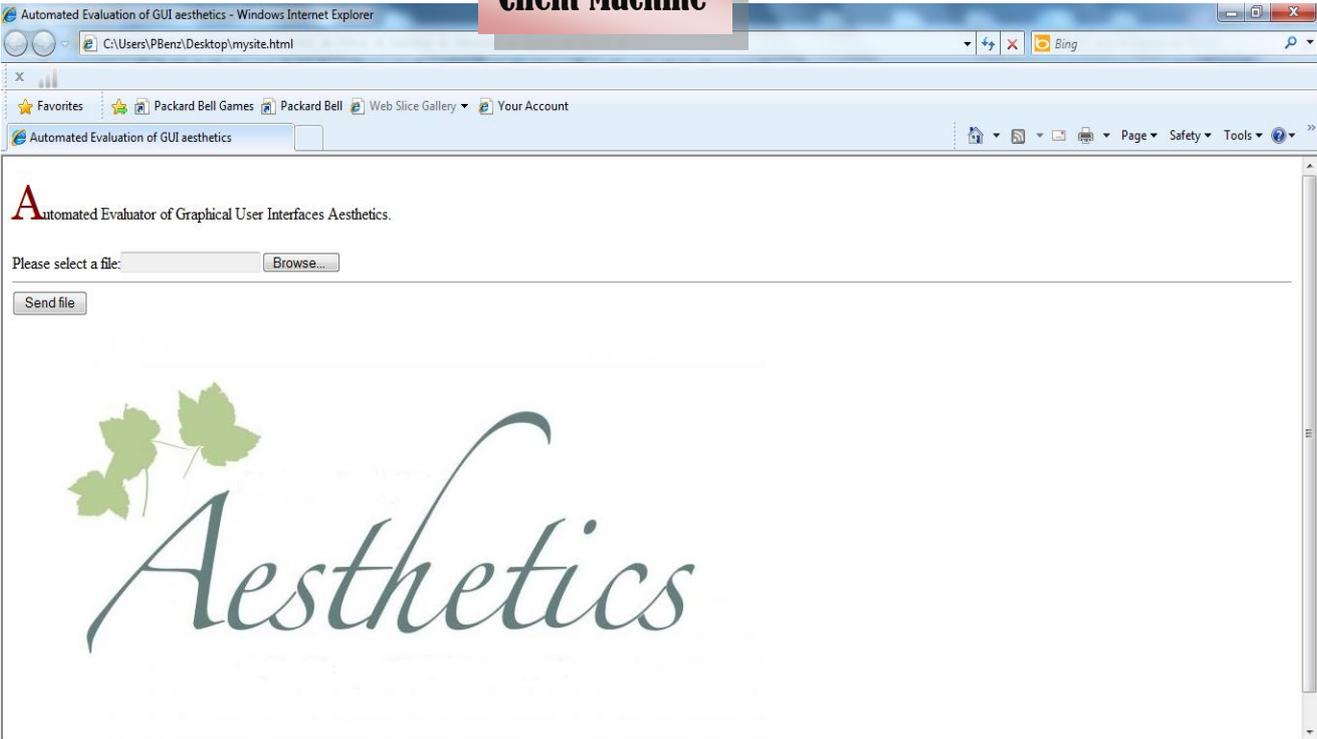
Origine ethnique

- Europe occidentale.
- Europe de l'est.
- Afrique du Nord.
- Afrique subsaharienne.
- Amérique du Nord.
- Amérique Latine.
- Proche orient
- Moyen orient
- Extreme orient
- Other:

Submit

Appendix E : Typical scenario with screenshots of actual Java code.

Client Machine



Automated Evaluator of Graphical User Interfaces Aesthetics.

Please select a file: Browse...



```
1 <html>
2 <head>
3 <title>Automated Evaluation of GUI aesthetics</title>
4 </head>
5 <body>
6 <p><font size="7" face="Georgia, Arial" color="maroon">A</font>utomated Evaluator
7 of Graphical User Interfaces Aesthetics.</p>
8 <form action="http://localhost:8200/AesthetiX/rest/upload" method="POST" enctype="multipart/form-data">
9 <div id="fileSection">
10 Please select a file:<input type="file" name="fileup0"/>
11 <!--File:<input type="file" name="fileup1" />
12 File:<input type="file" name="fileup2" />-->
13 </div>
14 <hr>
15 <input type="submit" value="Send file" />
16 <!--<input type="button" onclick="addFileElement()" value="Add File"/>-->
17 </form>
18 
19 </body>
20 </html>
21
```

Done Computer | Protected Mode: Off 100%

```

@Path("/upload")
public class FileUpload {
    private static File usiFile;
    private static String resultStatus = "<html>" + "<title>" + "Hello Jersey"
    + "</title>" + "<body><h1>" + "Failed!" + "</body></h1>"
    + "</html>";
    @POST
    @Produces(MediaType.TEXT_HTML)
    public String loadFileH(@Context HttpServletRequest request) {
        // this method uploads the file

        String fileRepository = "C:\\Users\\PBenz\\testRepo\\";

        if (ServletFileUpload.isMultipartContent(request)) {

            FileItemFactory factory = new DiskFileItemFactory();
            ServletFileUpload upload = new ServletFileUpload(factory);
            List<FileItem> items = null;
            try {
                items = upload.parseRequest(request);

            } catch (FileUploadException e) {

                e.printStackTrace();

            } catch (Exception e) {

                e.printStackTrace();

            }

            if (items != null) {

                Iterator<FileItem> iter = items.iterator();
                System.out.println(items.toString());
                while (iter.hasNext()) {

                    FileItem item = iter.next();
                    if (!item.isFormField() && item.getSize() > 0) {
                        String fileName = processFileName(item.getName());
                        try {
                            usiFile = new File(fileRepository + fileName);

                            item.write(usiFile);
                        } catch (Exception e) {
                            e.printStackTrace();
                        }
                        if (checkWellformedNess())
                            resultStatus = evaluate();
                    }

                }

            }

        }

        return resultStatus;
    }
}

```

2

3

5

```

public String evaluate(){
    String Result="" + "<title>" + "Hello Jersey" + "</title>"
    + "<body><h1>" + "Results" + "</h1>"+<table>";

    try {
        //parse the file
        DocumentBuilderFactory dbf = DocumentBuilderFactory.newInstance();
        DocumentBuilder db = dbf.newDocumentBuilder();
        Document doc = db.parse(usiFile);
        doc.getDocumentElement().normalize();
        /**System.out.println("Root element " + doc.getDocumentElement().getNodeName());*/
        NodeList nodeList = doc.getElementsByTagName("window");
        //evaluate the metrics on the file
        for (int s = 0; s < nodeList.getLength(); s++) {
            Node wndNode = nodeList.item(s);
            if (wndNode.getNodeType() == Node.ELEMENT_NODE)
            {
                Element wndElmnt = (Element)wndNode;
                float d = Metrics.density(wndElmnt);
                Result +=<tr><td>Density:</td><td>"+d+"</td></tr>";
                float bal = Metrics.balance(wndElmnt);
                Result +=<tr><td>Balance:</td><td>"+bal+"</td></tr>";
                float u = Metrics.unity(wndElmnt);
                Result +=<tr><td>Unity:</td><td>"+u+"</td></tr>";
                float sym = Metrics.symmetry(wndElmnt);
                Result +=<tr><td>Symmetry:</td><td>"+sym+"</td></tr>";
                float al = Metrics.alignment(wndElmnt);
                Result +=<tr><td>Alignment:</td><td>"+al+"</td></tr>";
                float gr = Metrics.grouping(wndElmnt);
                Result +=<tr><td>Grouping:</td><td>"+gr+"</td></tr>";
                double rep = Metrics.repartition(wndElmnt);
                Result +=<tr><td>Repartition:</td><td>"+rep+"</td></tr>";
                float sim = Metrics.simplicity(wndElmnt);
                Result +=<tr><td>Simplicity:</td><td>"+sim+"</td></tr>";
                float eco = Metrics.economy(wndElmnt);
                Result +=<tr><td>Economy:</td><td>"+eco+"</td></tr>";
                float p = Metrics.proportion(wndElmnt);
                Result +=<tr><td>Proportion:</td><td>"+p+"</td></tr>";
                double score = d*0.15 + bal*0.35 + u*0.10 + sym*0.10 + al*0.15 + gr*0.03 + rep*0.03 + sim*0.03 + eco*0.03 + p*0.03;
                Result +=<tr><td>Score(weighted sum):</td><td>"+score+"</td></tr>";
            }
        }
        Result +=</table>"+<br>+"Done! (if you cannot see any results please check that the file you uploaded is a valid UsiXML fi
    }
    catch (Exception e){
        e.printStackTrace();
    }

    return Result;
}

```



```

package aesthetiX;

import org.w3c.dom.Element;

public class Metrics {
//-----
/**name:DENSITY
**importance: low
**ref:[2]
**formula:1-[(sum Ai)/ Aframe]*//
//-----/
    public static float density(Element elmnt)// counts the widgets in the interface
    {
        int areaFrame = 0;
        int sumOfAreas = 0;
        NodeList nodeList = elmnt.getElementsByTagName("gridBagBox");
        for (int s = 0; s < nodeList.getLength(); s++)
        {
            //computer the total area of the frame
            int totalw = Integer.parseInt(((Element) nodeList.item(s)).getAttribute("gridWidth"));
            int totalh = Integer.parseInt(((Element) nodeList.item(s)).getAttribute("gridHeight"))-1;
            areaFrame = totalw * totalh;
            //System.out.println("area total=" + areaFrame);

            //scan the widgets of the frame one at a time
            NodeList kids = nodeList.item(s).getChildNodes();
            for (int i = 0; i < kids.getLength(); i++)
            {
                if (kids.item(i).getNodeName() != "#text")
                {
                    //System.out.println("child node n°" + i + " :"+ kids.item(i).getNodeName());
                    //compute area of widget and add to sum
                    int width = Integer.parseInt(((Element) kids.item(i)).getAttribute("gridwidth"));
                    int height = Integer.parseInt(((Element) kids.item(i)).getAttribute("gridheight"));
                    sumOfAreas = sumOfAreas + (width * height);
                }
            }
            //System.out.println("somme des areas=" + sumOfAreas);
        }
        return 1 - ((float) sumOfAreas / (float) areaFrame);
    }
//-----
/**name:BALANCE
**importance: high
**ref:[5]
**formula:1-[(|BMvertical| + |BMhorizontal|)/2]*//
//-----/
    public static float balance(Element elmnt)
    {
        float Wl = 0;
        float Wr = 0;
        float Wt = 0;
        float Wb = 0;
        float BMvert = 0;

```

Client Machine

Results

Density:	0.4100877
Balance:	0.6110871
Unity:	0.26666665
Symmetry:	0.7713198
Alignment:	0.84880954
Grouping:	0.6477578
Repartition:	0.0075
Simplicity:	0.069767445
Economy:	0.125
Proportion:	0.63293254
Score(weighted sum):	0.5510024463295938

Done! (if you cannot see any results please check that the file you uploaded is a valid UsiXML file)